

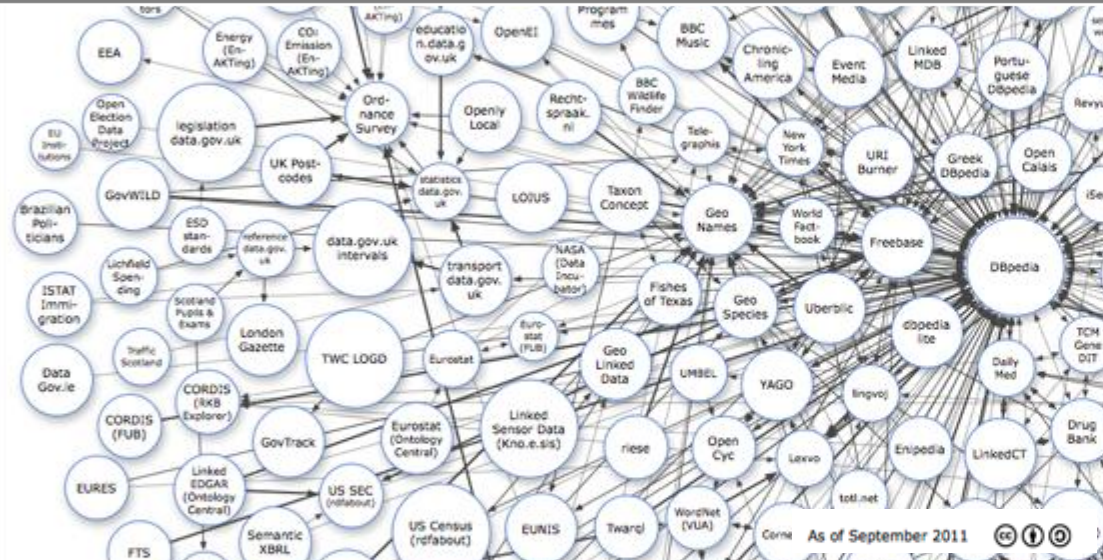
Accepting the XBRL Challenge with Linked Data for Financial Data Integration

Benedikt Kämpgen, Tobias Weller, Sean O’Riain, Craig Weber, Andreas Harth

11th ESWC 2014 In-Use

27 May 2014

Institute of Applied Informatics and Formal Description Methods (AIFB)



Mechanics in Financial Analysis



EDINET - Mozilla Firefox

Semantic Stati... GESIS: DDI Schloss Dagstu... DDI and Sema... Edit Investigati... Home Page for... Financial Infor... EDINET

info.edinet-fsa.go.jp

EDINET
Electronic Disclosure for Investors' NETwork

EDINET 金融商品取引法に基づく有価証券報告書等の開示書類に関する電子開示システム

有価証券報告書等の開示書類を閲覧するホームページ

お知らせ
推奨端末仕様
操作ガイド
よくある質問(Q&A)
当サイトに関するご意見・ご感想

はじめに

EDINET (Electronic Disclosure for Investors' NETwork) は、『金融商品取引法に基づく有価証券報告書等の開示書類に関する電子開示システム』のことで、提出された開示書類について、インターネット上においても閲覧を可能とするものです。

EDINETにアクセスした際に、「セキュリティ証明書」についての画面が表示された場合は、端末への設定が必要となります。端末設定方法に関する詳細、およびブラウザなどの推奨環境については、[推奨端末仕様](#)をご覧ください。

その他、EDINETについての詳細は、[操作ガイド](#)の「EDINET概要書」および各種ガイドをご覧ください。

よくある質問(Q&A)が2012年6月29日に更新されていますのでご覧ください。

システムメンテナンスに伴うサービス一時停止のお知らせ

以下の日時、サービスが停止しますのでお知らせします。
平成 24年 10月 13日(土) 0:00 ~ 平成 24年 10月 14日(日) 7:30
利用者の皆様にはご協力のほど、よろしくお願いいたします。

注意(未公開株、ファンド等の取引について)

●「未公開株」「私募債」「ファンド(組合など)」の証券の取引に関してトラブルが発生しており、また、「EDINETに証券の情報を開示しているので安心である。」「財務局受理印が押されているので安心である。」といった違法的な勧誘が行われているとの情報が寄せられています。

●金融庁・財務局はEDINETに開示している情報について、
①開示内容について真実・正確と認定すること
②開示内容に投資判断上の重要な事項の記載が欠けていないと認定すること
③開示している有価証券の価値を保証・承認すること
は一切ありませんので、くれぐれもご注意ください。

●なお、有価証券の取引に関し、上記①～③のような表示を行うことは、法により禁止されておりますので、ご注意ください。

お知らせ

Task: Find value of Assets of a company 三精輸送機

Mechanics in Financial Analysis



EDINET - Mozilla Firefox

Semantic Stati... GESIS: DDI LZI Schloss Dagstu... DDI and Seman... Edit Investigati... Home Page for... Financial Infor... EDINET

info.edinet-fsa.go.jp

ださい。

お知らせ
○EDINETの受理時間について
株券の発行において、当日の市場価格の終値を参照して発行価格が確定し、当日又は翌日に効力を生じさせることが必要な訂正届出書について、発行価格の協議に時間を要する場合、書類作成に時間を要する場合、システムの都合等やむを得ない理由があるときは、原則として、EDINETの受付時間を延長することになりました（おおむね19時まで）。※なお、その他の書類の開示についても、EDINETの受付時間を延長することがあります。

更新情報
平成24年10月4日 [次世代EDINETタクソノミ\(案\)第二版並びに提出者向け事前チェックテスト実施要領及び参加要領](#)が公表されました。
平成24年10月4日 [提出書類ファイル仕様書](#)が更新されました。
平成24年6月29日 [よくある質問](#)が更新されました。(情報公開)
平成24年5月25日 [企業別タクソノミ作成ガイドライン\(その2:IFRS適用提出者用\)](#)が更新されました。
平成24年4月11日 [推奨端末仕様](#)が更新されました。
平成24年3月14日 [2012年版EDINETタクソノミ及びXBRL作成ガイド](#)が公開されました。
平成24年3月14日 [EDINET概要書](#)が更新されました。
平成24年1月28日 [大量保有報告書様式\(EXCEL版\)の使い方について](#)が更新されました。
平成23年7月25日 [2011年版EDINETタクソノミの追加タクソノミ](#)が公開されました。
平成23年3月14日 [2011年版EDINETタクソノミ及びXBRL作成ガイド](#)が公開されました。

閲覧

→ 有価証券報告書等 → 公 告

ダウンロード

→ XBRL一括 → EDINETタクソノミ → EDINETコードリスト
→ 提出書類様式 → EDINETコード対応表

金融庁/Financial Services Agency, The Japanese Government Copyright©金融庁 All Rights Reserved.

Example Craig Weber

Task: Find value of Assets of a company 三精輸送機

Mechanics in Financial Analysis



EDINET - Mozilla Firefox

info.edinet-fsa.go.jp

ださい。

お知らせ
○EDINETの受理時間について
株券の発行において、当日の市場価格の終値を参照して発行価格が確定し、当日又は翌日に効力を生じさせることが必要な訂正届出書について、発行価格の協議に時間を要する場合、書類作成に時間を要する場合、システムの都合等やむを得ない理由があるときは、原則として、EDINETの受付時間を延長することになりました（おおむね19時まで）。※なお、その他の書類の開示についても、EDINETの受付時間を延長することがあります。

更新情報
平成24年10月4日 [次世代EDINETタクソノミ\(案\)第二版並びに提出者向け事前チェックテスト実施要領及び参加要領](#)が公表されました。
平成24年10月4日 [提出書類ファイル仕様書](#)が更新されました。
平成24年6月29日 [よくある質問](#)が更新されました。(情報公開)
平成24年5月25日 [企業別タクソノミ作成ガイドライン\(その2:IFRS適用提出者用\)](#)が更新されました。
平成24年4月11日 [推奨端末仕様](#)が更新されました。
平成24年3月14日 [2012年版EDINETタクソノミ及びXBRL作成ガイド](#)が公開されました。
平成24年3月14日 [EDINET概要書](#)が更新されました。
平成24年1月28日 [大量保有報告書様式\(EXCEL版\)の使い方について](#)が更新されました。
平成23年7月25日 [2011年版EDINETタクソノミの追加タクソノミ](#)が公開されました。
平成23年3月14日 [2011年版EDINETタクソノミ及びXBRL作成ガイド](#)が公開されました。

関 覧

有価証券報告書等

公告

ダウンロード

[XBRL一括](#) [EDINETタクソノミ](#) [EDINETコードリスト](#)


[提出書類様式](#) [EDINETコード対応表](#)

金融庁/Financial Services Agency, The Japanese Government Copyright©金融庁 All Rights Reserved.

Example Craig Weber

Extensible Business Reporting Language (XBRL): Describes business reporting information in an XML format, for exchange, and comparison [XBRL SPEC].

XBRL – Example



View Filing Data

SEC Home » Search the Next-Generation EDGAR System » Company Search » Current Page

MASTERCARD INC (Filer) CIK: 0001141391


Print Document View Excel Document

Cover	Consolidated Balance Sheet (USD \$)	Mar. 31, 2014	Dec. 31, 2013
Document and Entity Information	In Millions, unless otherwise specified		
Financial Statements	ASSETS		
Consolidated Balance Sheet	Cash and cash equivalents	\$ 4,019	\$ 3,599
Consolidated Balance Sheet (Parenthetical)	Restricted cash for litigation settlement	540	723
Consolidated Statement of Operations	Investment securities available-for-sale, at fair value	2,543	2,696
Consolidated Statement of Comprehensive Income	Accounts receivable	995	966
Consolidated Statement of Changes in Equity	Settlement due from customers	1,476	1,351
	Restricted security deposits held for customers	913	911
	Prepaid expenses and other current assets	552	471
	Deferred income taxes	304	233
	Total Current Assets	11,342	10,950
	Property, plant and equipment, net of		

- Mastercard
- 11.34B USD
- Total Current Assets
- Q1 2014

[http://www.sec.gov/edgar/
searchedgar/webusers.htm](http://www.sec.gov/edgar/searchedgar/webusers.htm)

XBRL – Example



View Filing Data

SEC Home » Search the Next-Generation EDGAR System » Company Search » Current Page

MASTERCARD INC (Filer) CIK: 0001141391


[Print Document](#) [View Excel Document](#)

Cover	Consolidated Balance Sheet (USD \$)	Mar. 31, 2014	Dec. 31, 2013
Document and Entity Information	In Millions, unless otherwise specified		
Financial Statements	ASSETS		
Consolidated Balance Sheet	Cash and cash equivalents	\$ 4,019	\$ 3,599
Consolidated Balance Sheet (Parenthetical)	Restricted cash for litigation settlement	540	723
Consolidated Statement of Operations	Investment securities available-for-sale, at fair value	2,543	2,696
Consolidated Statement of Comprehensive Income	Accounts receivable	995	966
Consolidated Statement of Changes in Equity	Settlement due from customers	1,476	1,351
	Restricted security deposits held for customers	913	911
	Prepaid expenses and other current assets	552	471
	Deferred income taxes	304	233
	Total Current Assets	11,342	10,950
	Property, plant and equipment, net of		

- Mastercard
- 11.34B USD
- Total Current Assets
- Q1 2014

[http://www.sec.gov/edgar/
searchedgar/webusers.htm](http://www.sec.gov/edgar/searchedgar/webusers.htm)

XBRL – Example



View Filing Data

SEC Home » Search the Next-Generation EDGAR System » Company Search » Current Page

MASTERCARD INC (Filer) CIK: 0001141391

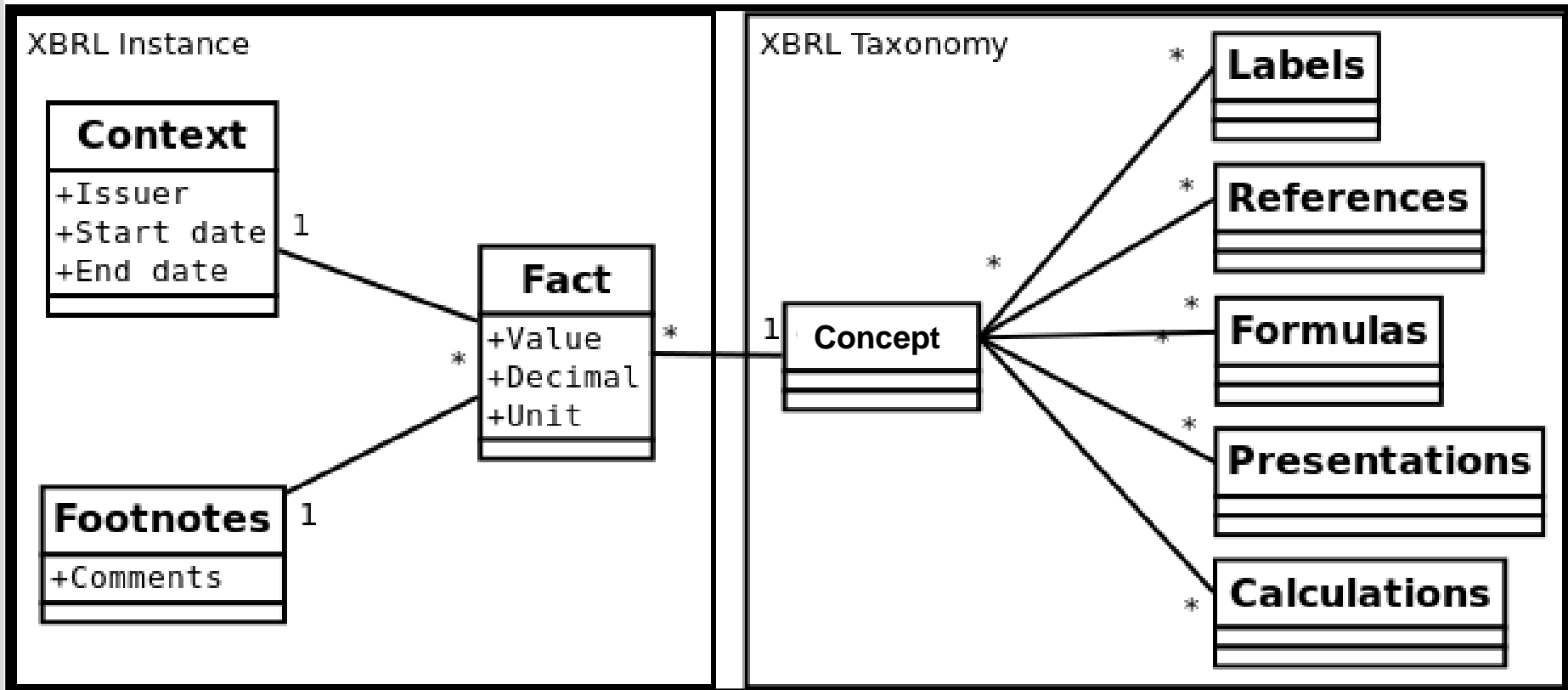
Print Document View Excel Document

Cover	Consolidated Balance Sheet (USD \$) In Millions, unless otherwise specified	Mar. 31, 2014	Dec. 31, 2013
Document and Entity Information	ASSETS		
Financial Statements	Cash and cash equivalents	\$ 4,019	\$ 3,599
Consolidated Balance Sheet	Restricted cash for litigation settlement	540	723
Consolidated Balance Sheet (Parenthetical)	Investment securities available-for-sale, at fair value	2,543	2,696
Consolidated Statement of Operations	Accounts receivable	995	966
Consolidated Statement of Comprehensive Income	Settlement due from customers	1,476	1,351
Consolidated Statement of Changes in Equity	Restricted security deposits held for customers	913	911
	Prepaid expenses and other current assets	552	471
	Deferred income taxes	304	233
	Total Current Assets	11,342	10,950
	Property, plant and equipment, net of		

- Mastercard
 - 11.34B USD
 - Total Current Assets
 - Q1 2014
- U.S. Securities and Exchange publishes balance sheets of 8,000 U.S. companies

<http://www.sec.gov/edgar/searchedgar/webusers.htm>

XBRL – Data Model



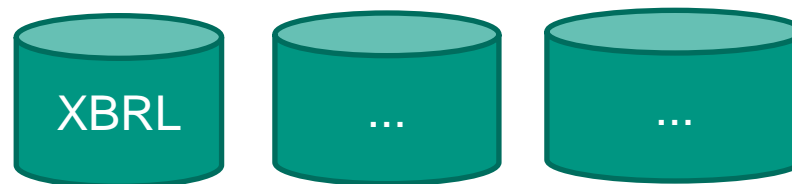
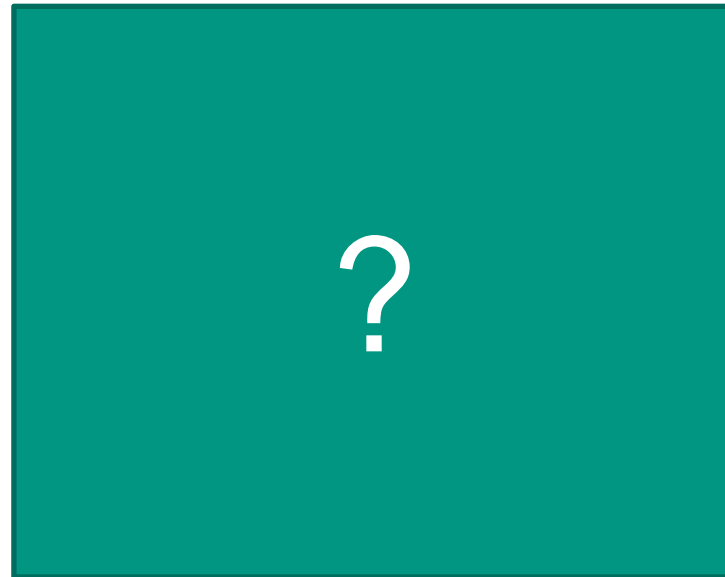
Simplified structure of XBRL

Outline

- Mechanics in Financial Analysis
- **XBRL – Challenges to Data Integration**
- Financial Information Observation System (FIOS)
- FIOS in the XBRL Challenge
- Benefits of Semantics
- FIOS Competitors

XBRL – Challenges to Data Integration

Analysts



Data Sources

XBRL – Challenges to Data Integration – (1)

Analysts

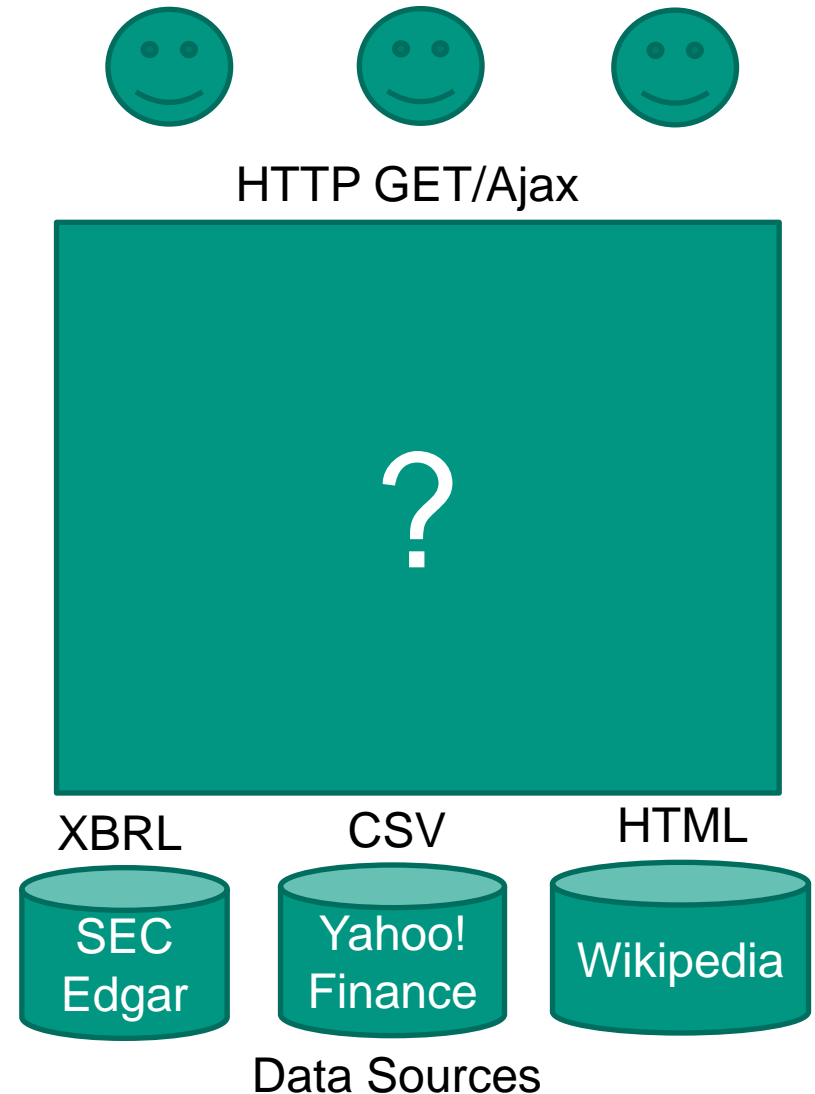
Problem:

No globally accepted
schema for finance data

Example:

Balance sheets from SEC
Stock Market values from
Yahoo! Finance

Number of employees from
Wikipedia



XBRL – Challenges to Data Integration – (2)

Analysts

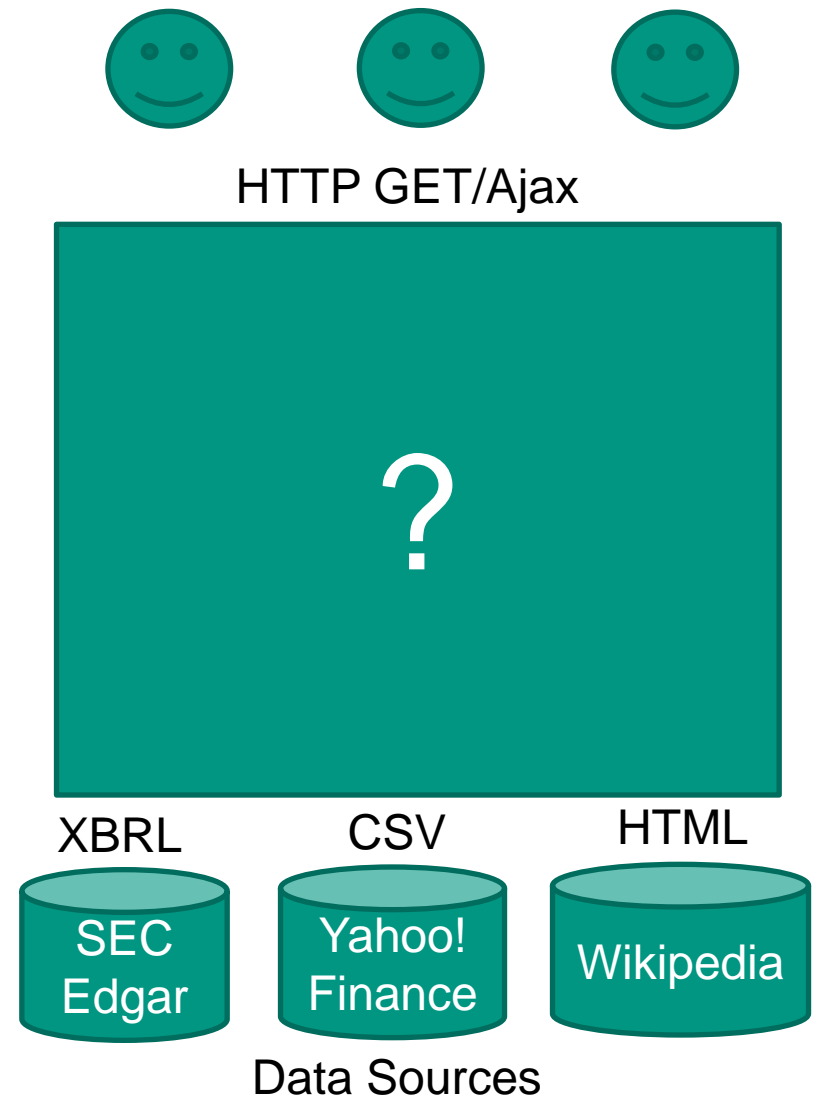
Problem:

Different taxonomies per
branch, country, version

No alignment of taxonomies

Example:

Total Assets from
US-GAAP 2009, 2011



XBRL – Challenges to Data Integration – (3)

Analysts

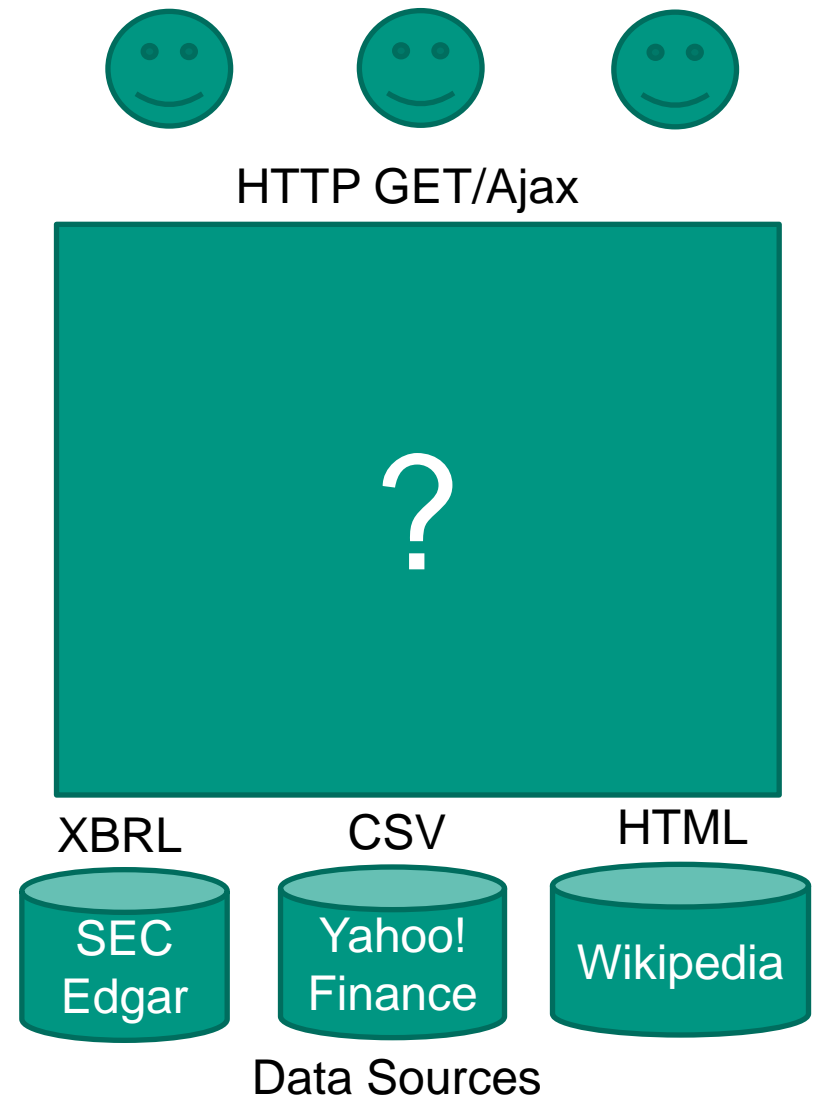
Problem:

Missing unique identifiers
for companies

Example:

Earnings per Share vs.
Price per Share

1141391 / MA / Mastercard

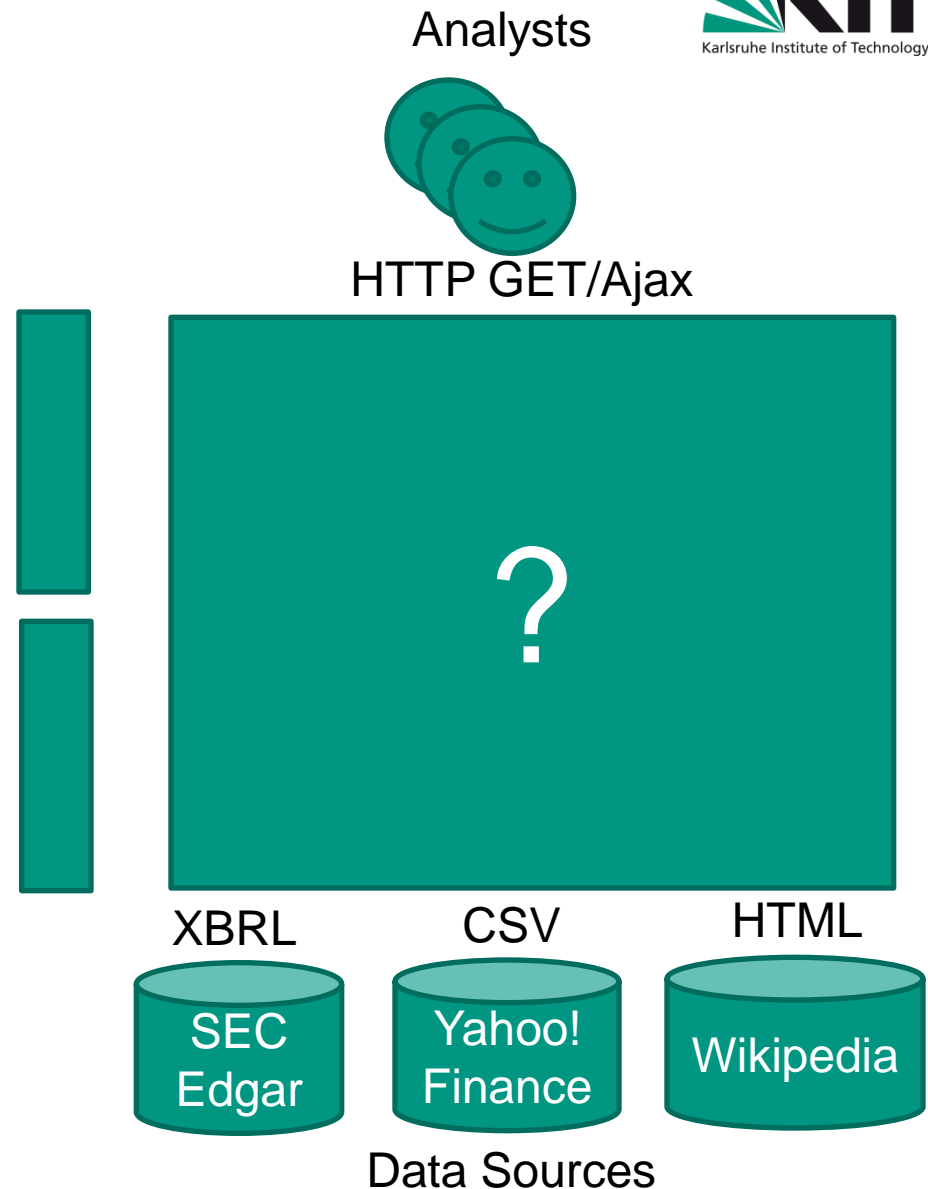


Outline

- Mechanics in Financial Analysis
- XBRL – Challenges to Data Integration
- **Financial Information Observation System (FIOS)**
- FIOS in the XBRL Challenge
- Benefits of Semantics
- FIOS Competitors

Financial Information Observation System (FIOS)

- Online Analysis components
- Offline ETL components



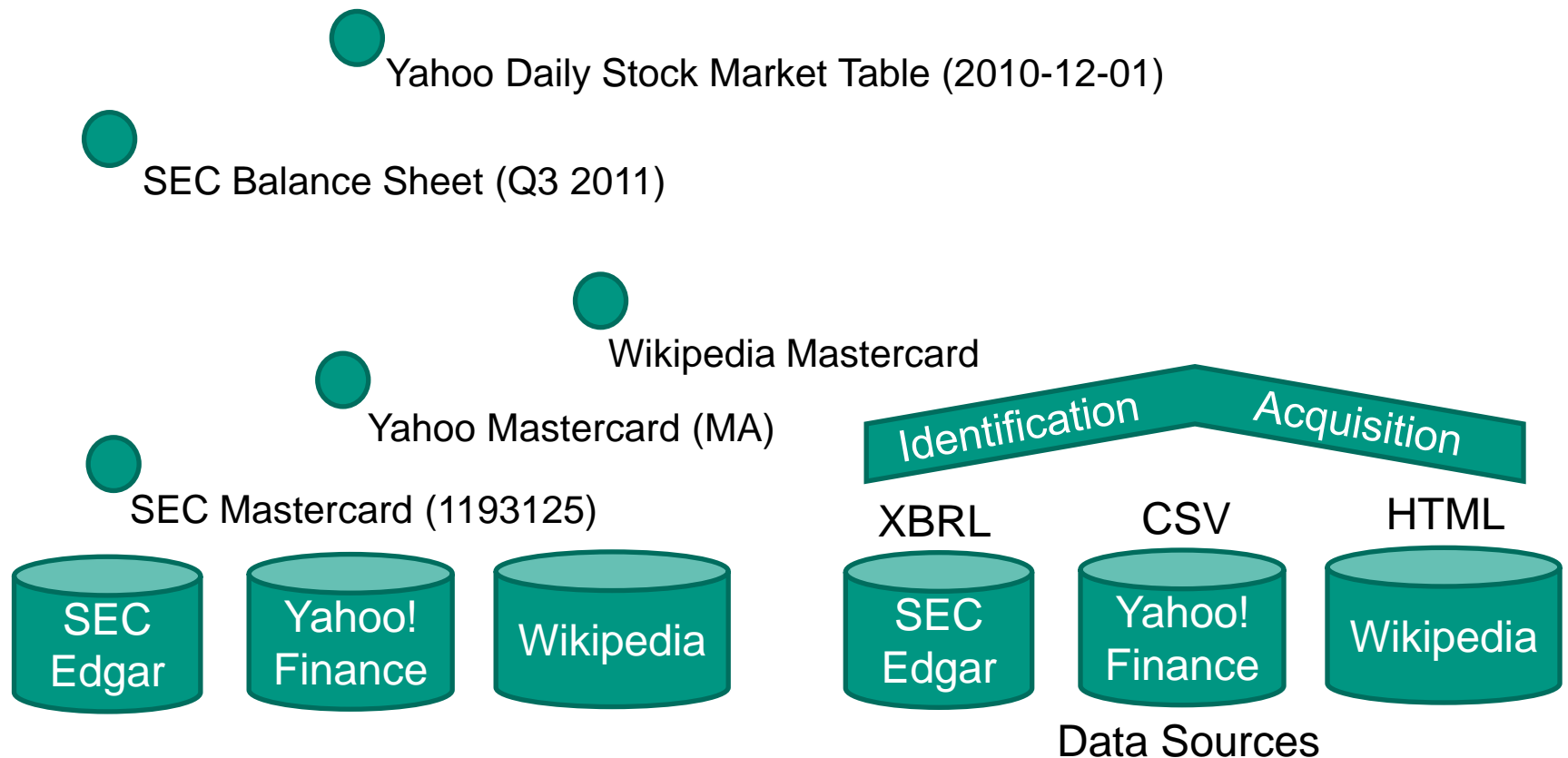
fios – Gaelic „Knowledge“

FIOS – How to find relevant information?

Analysts

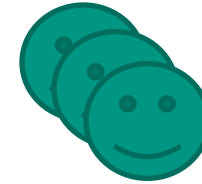


HTTP GET/Ajax

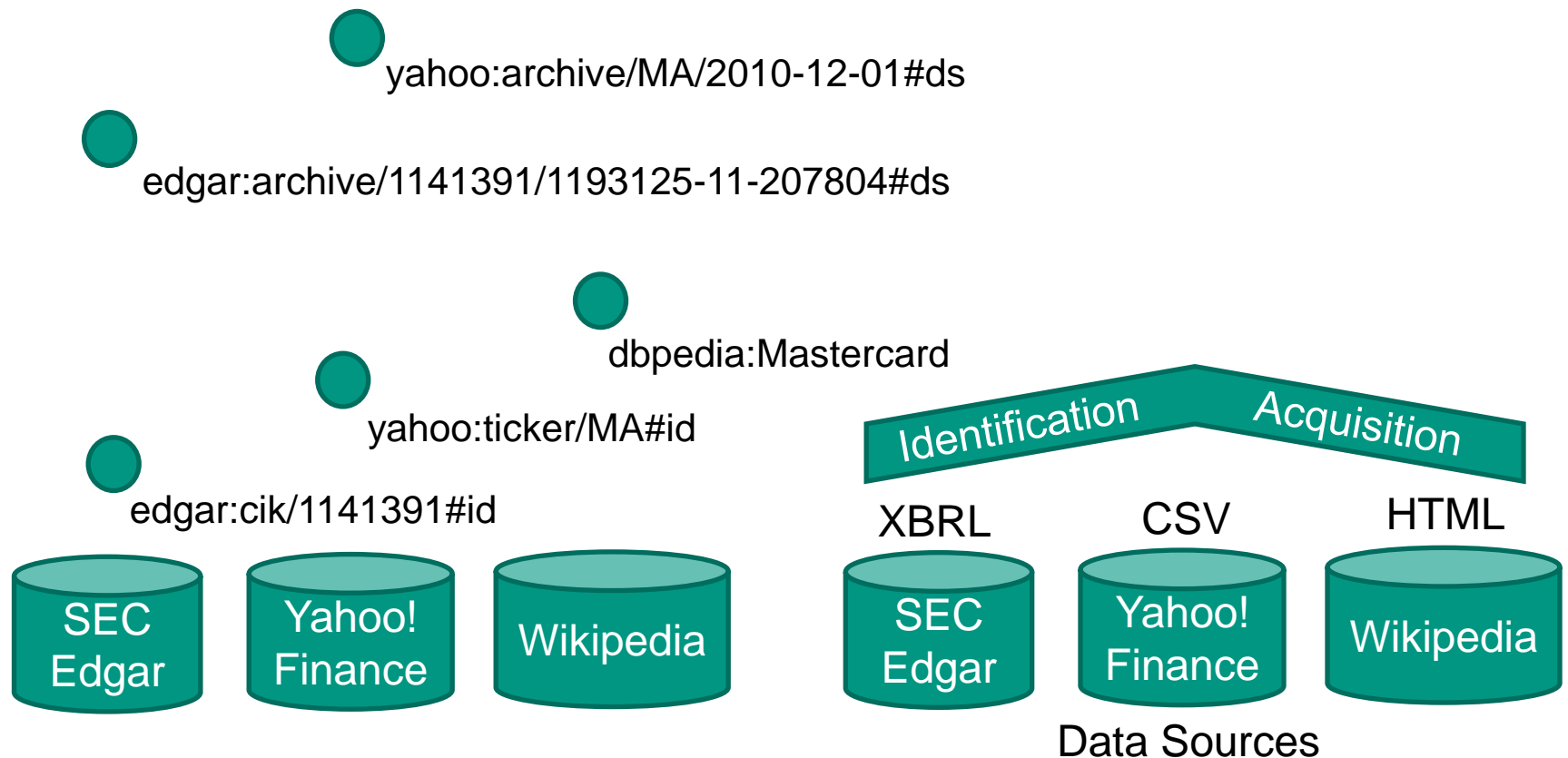


FIOS – How to find relevant information?

Analysts

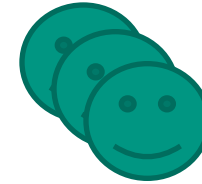


HTTP GET/Ajax

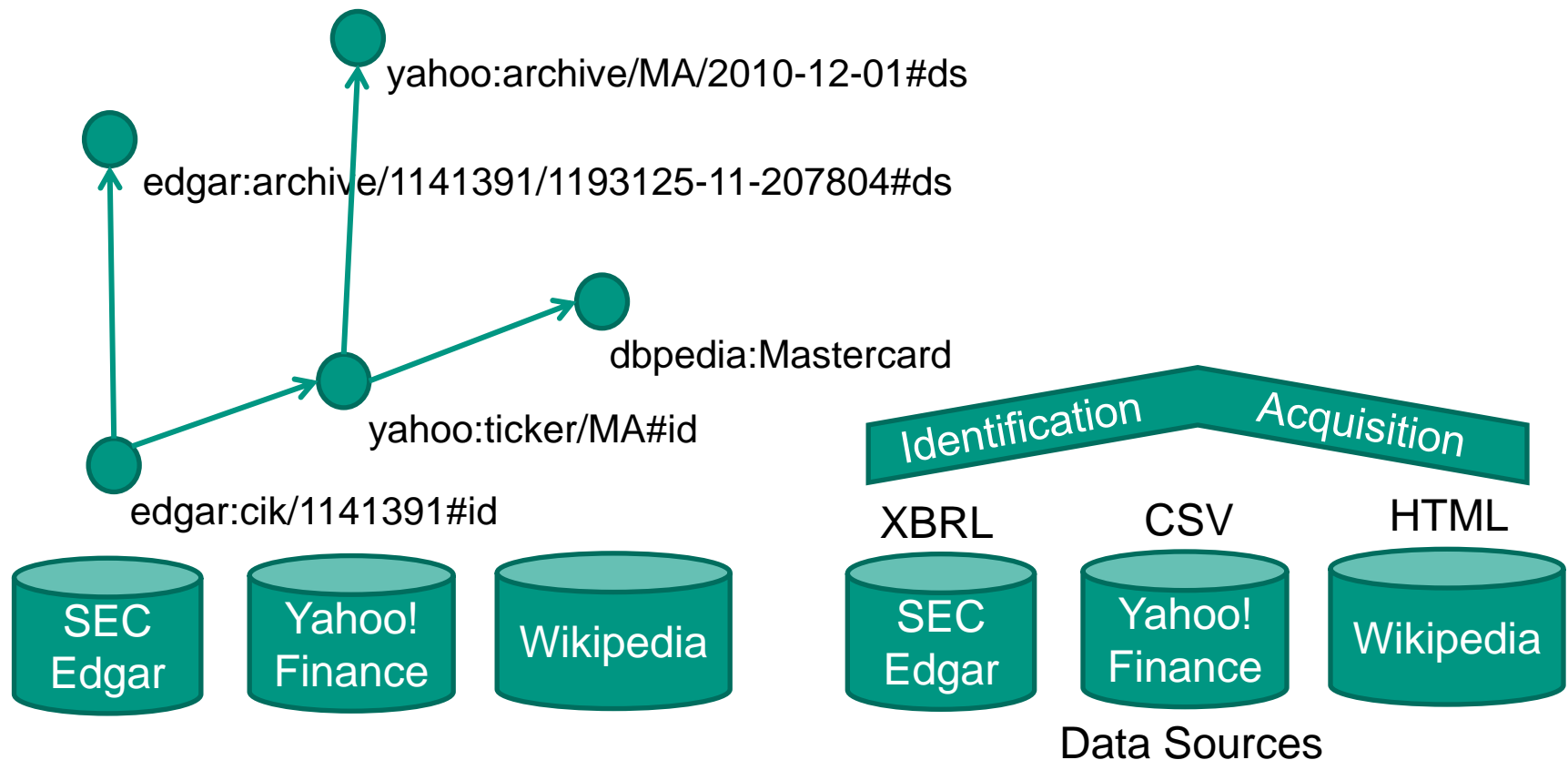


FIOS – How to find relevant information?

Analysts



HTTP GET/Ajax

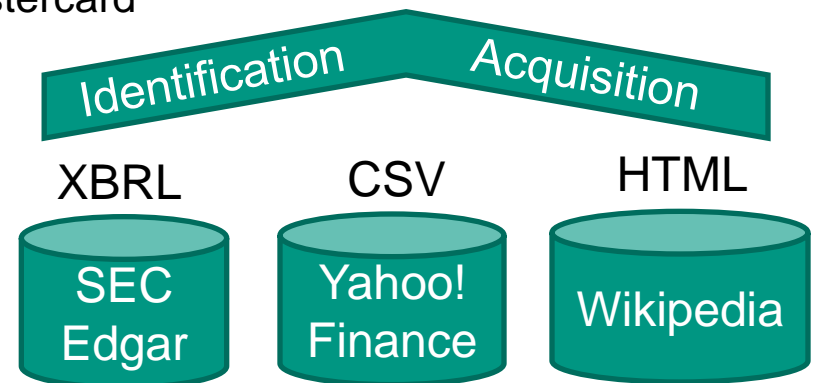
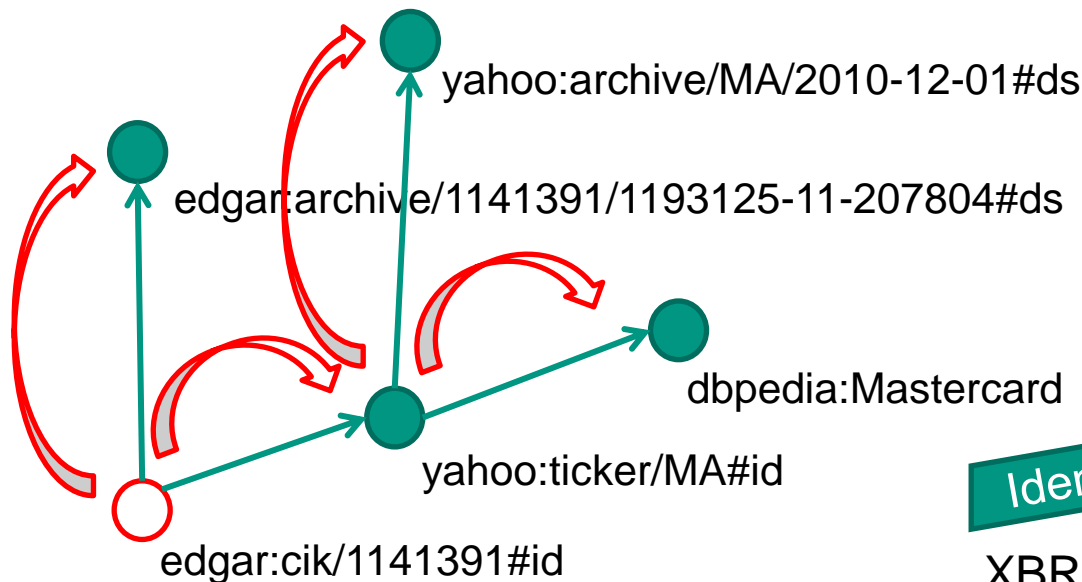


FIOS – How to find relevant information?

Analysts

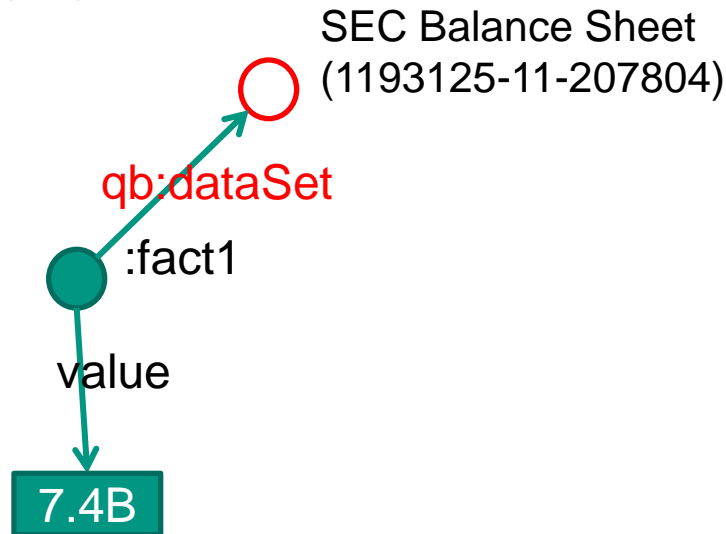



HTTP GET/Ajax

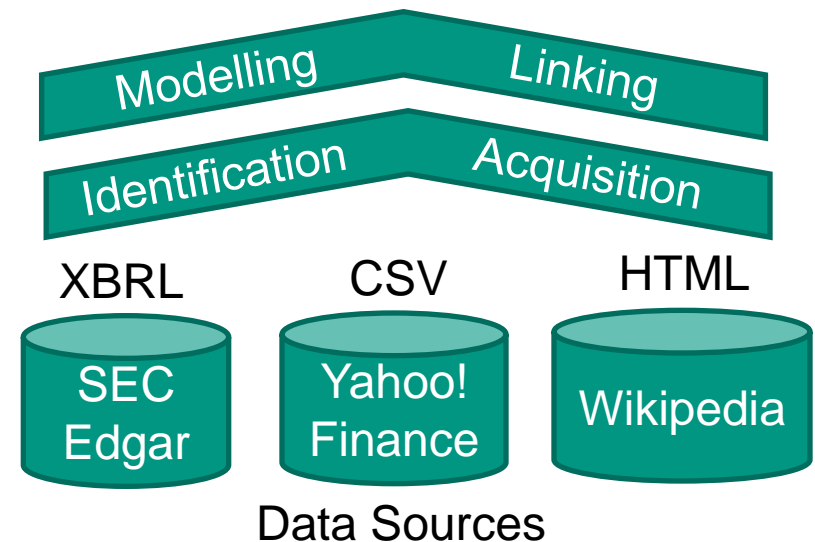


Data Sources

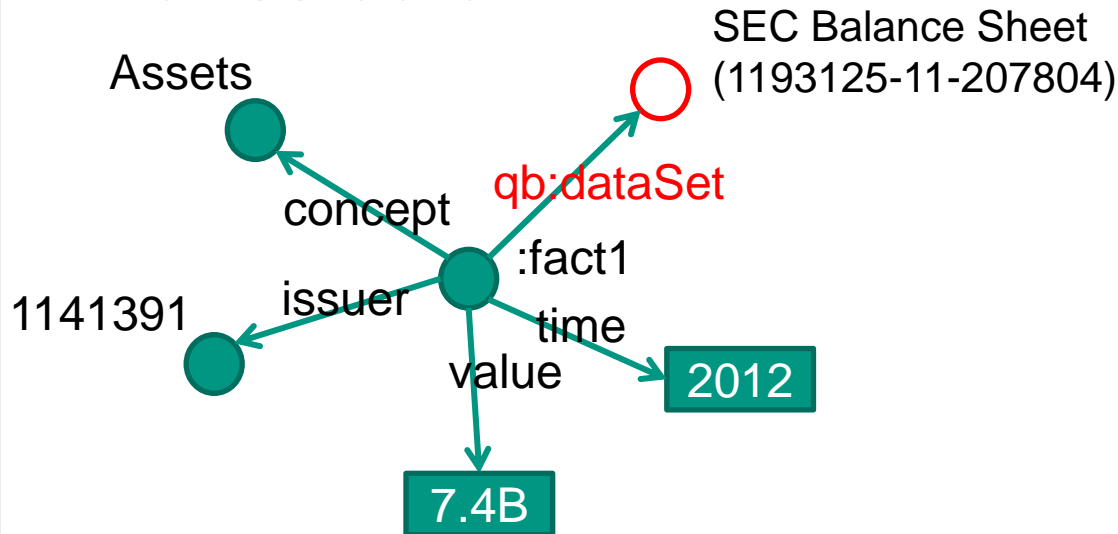
FIOS – How to describe finance data?



Analysts

HTTP GET/Ajax



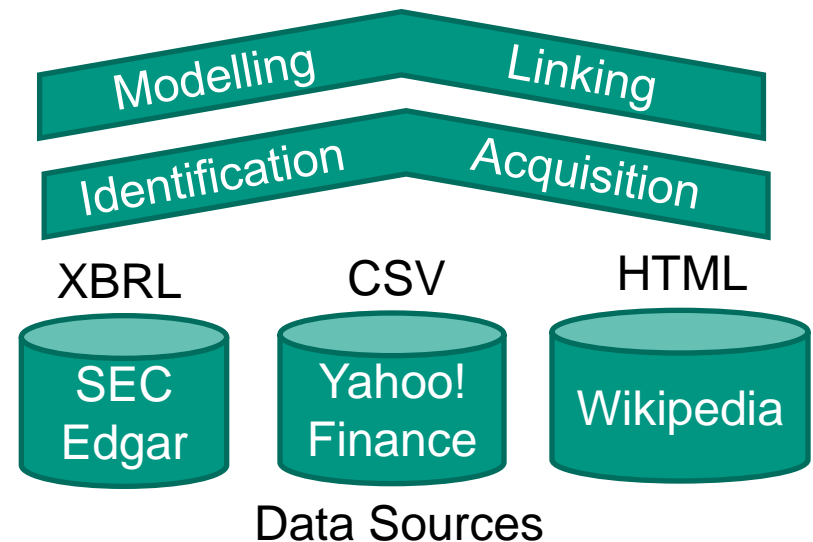
FIOS – How to describe finance data?



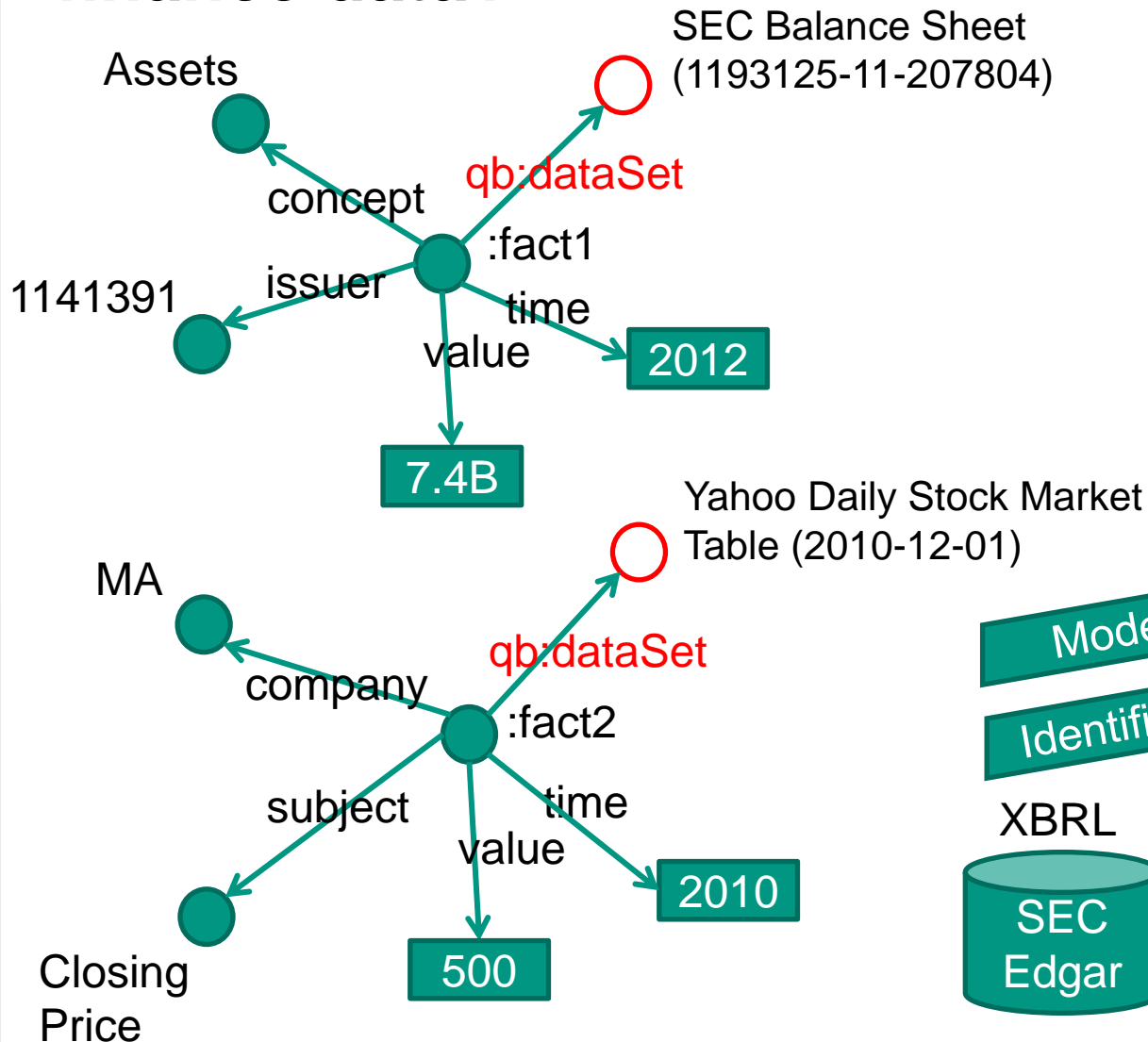
Analysts




HTTP GET/Ajax



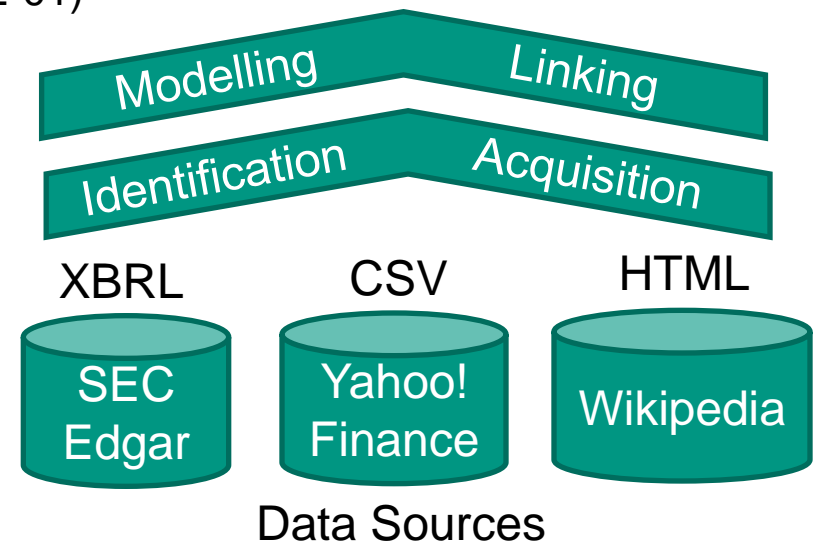
FIOS – How to describe finance data?



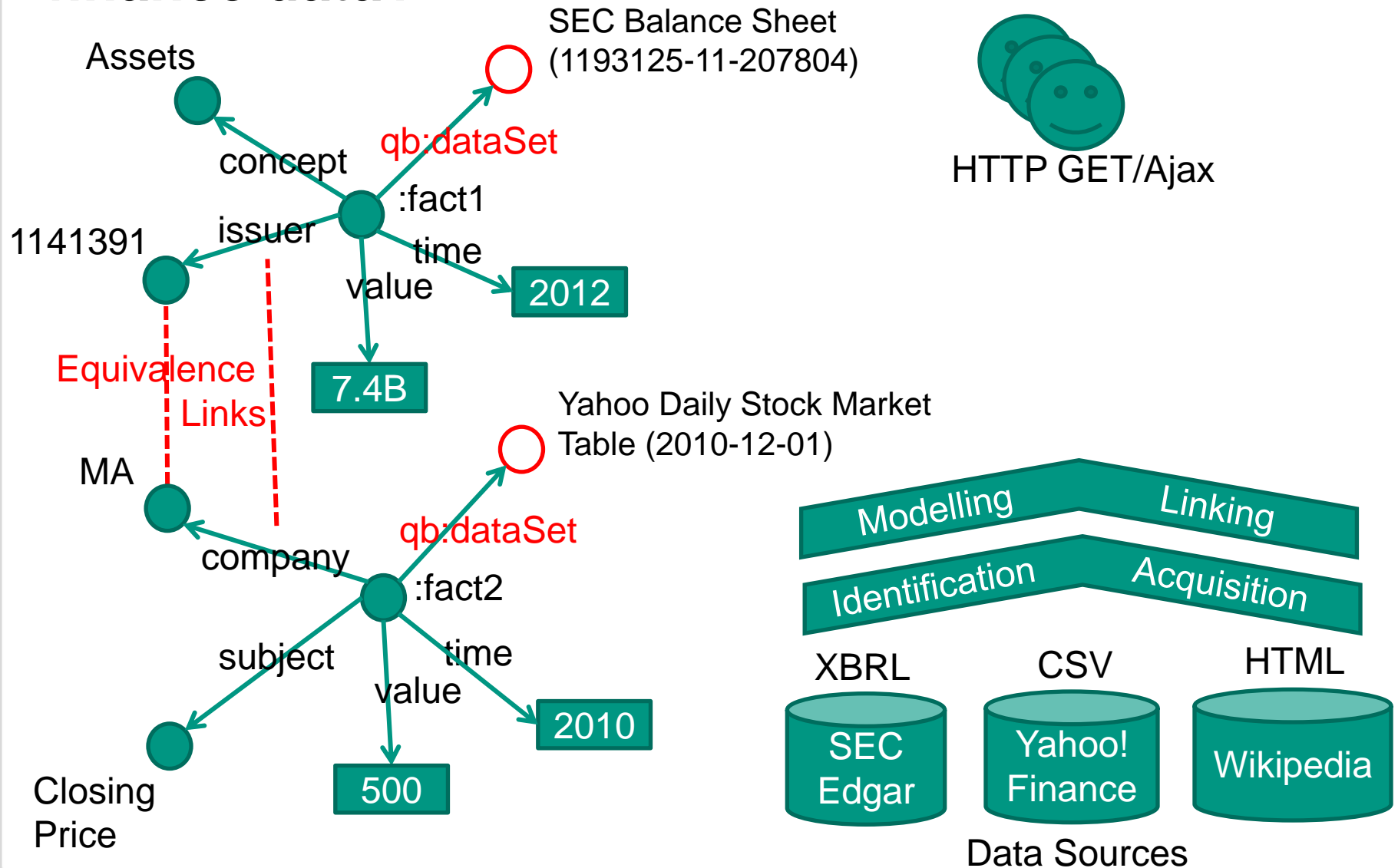
Analysts



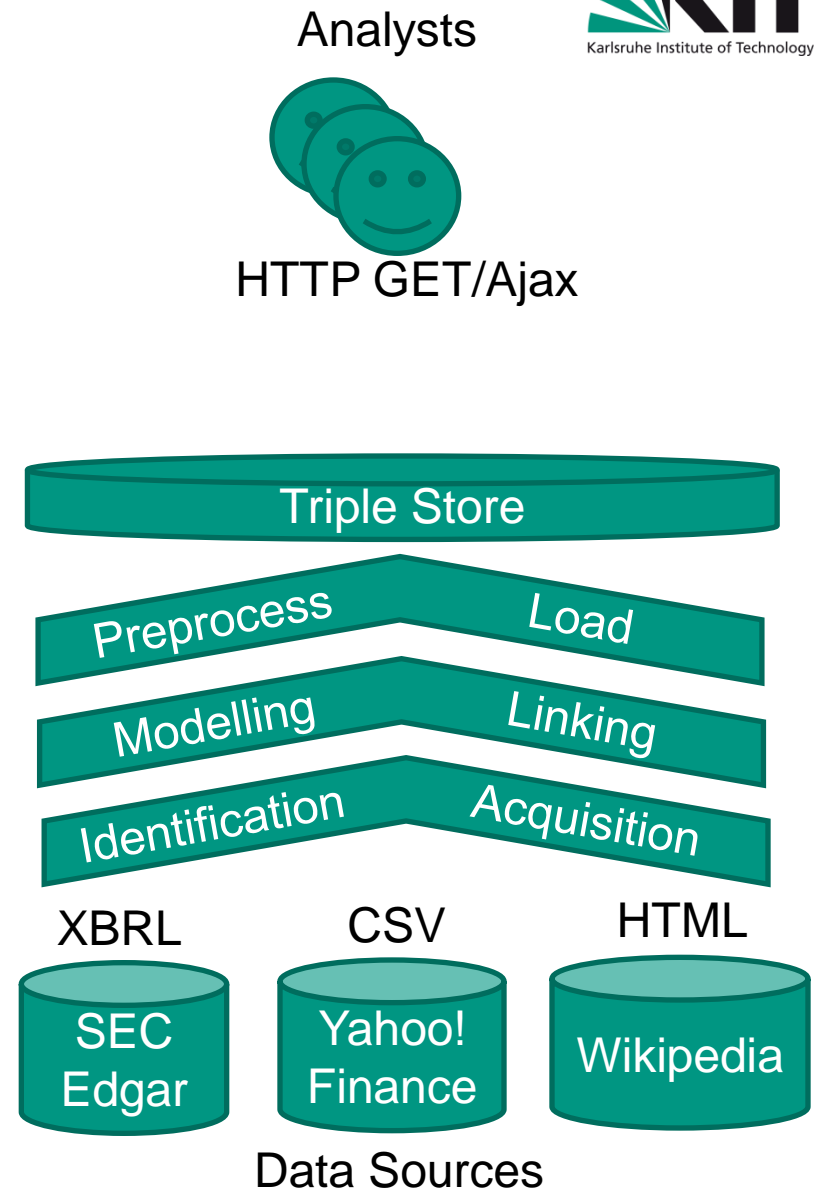
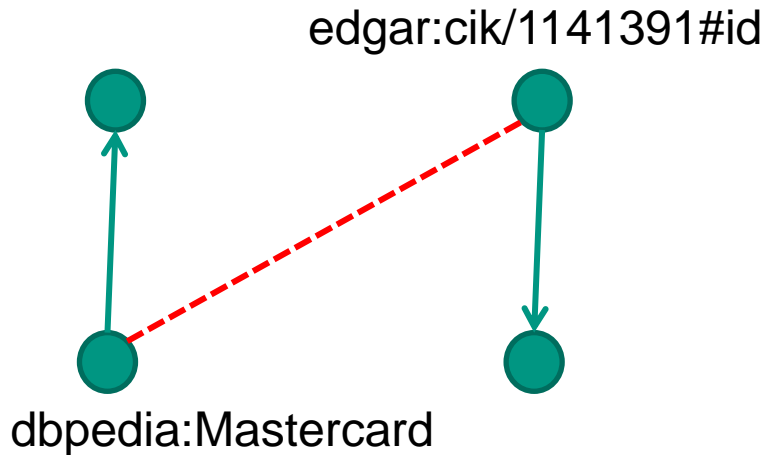
HTTP GET/Ajax



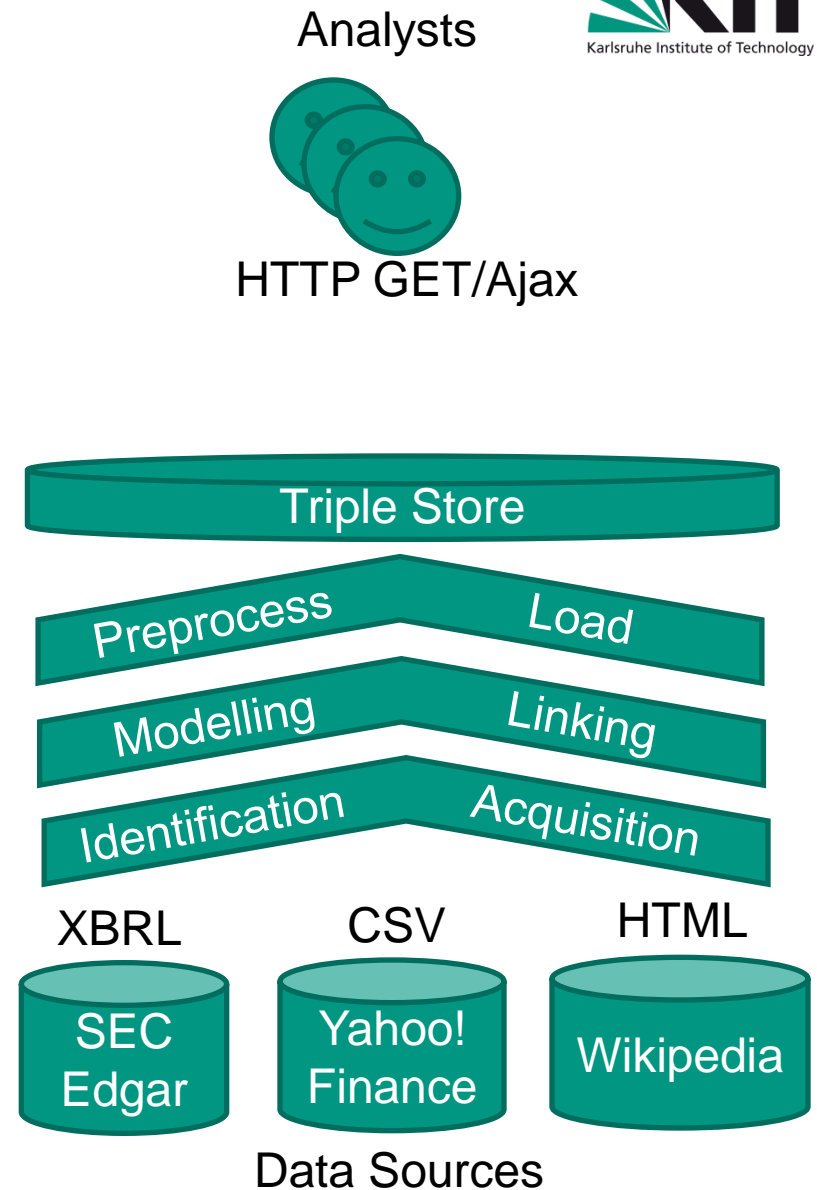
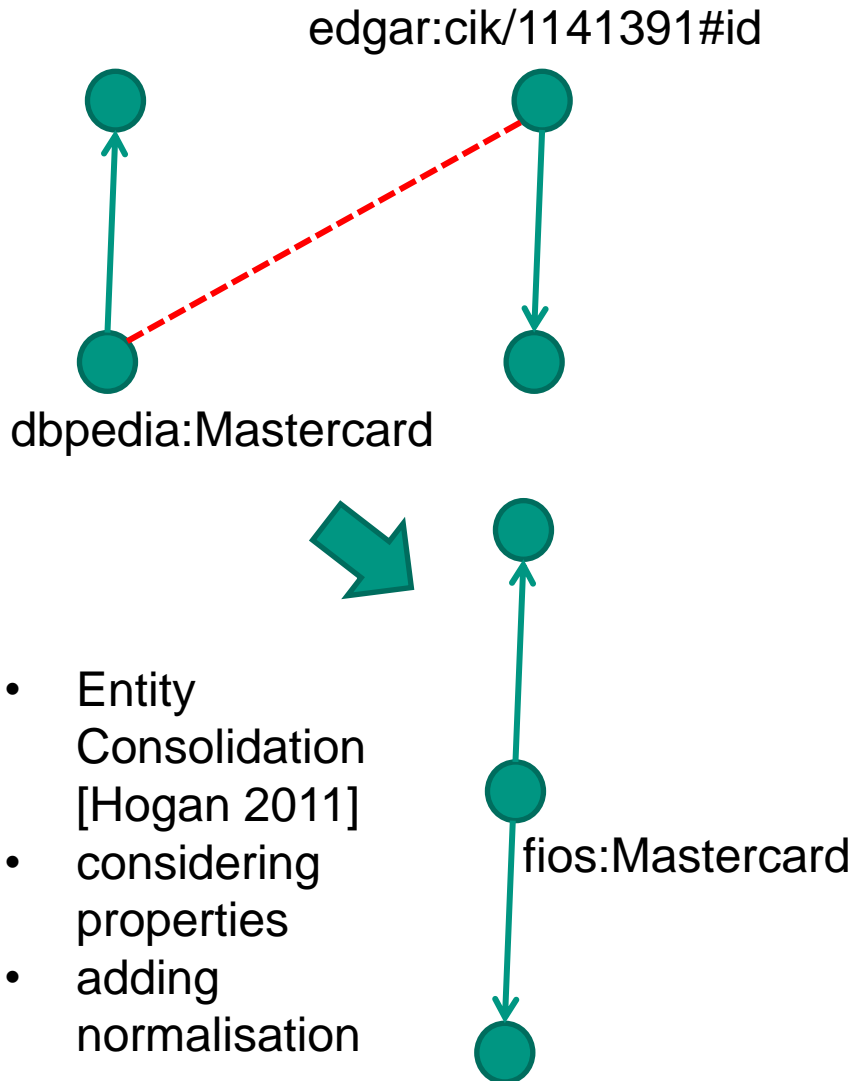
FIOS – How to describe finance data?



FIOS – How to consider equivalence?

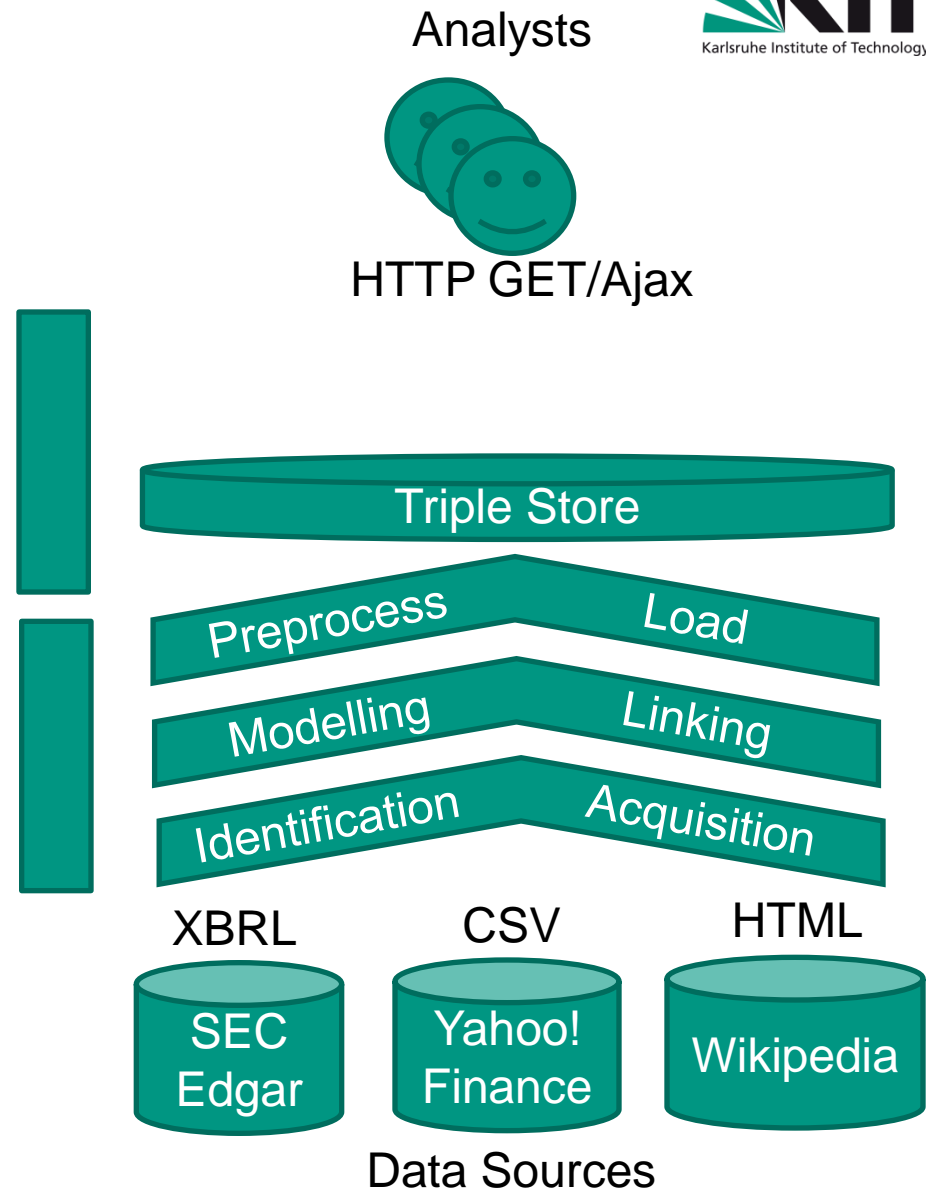


FIOS – How to consider equivalence?



Financial Information Observation System (FIOS)

- Online Analysis components
- Offline ETL components

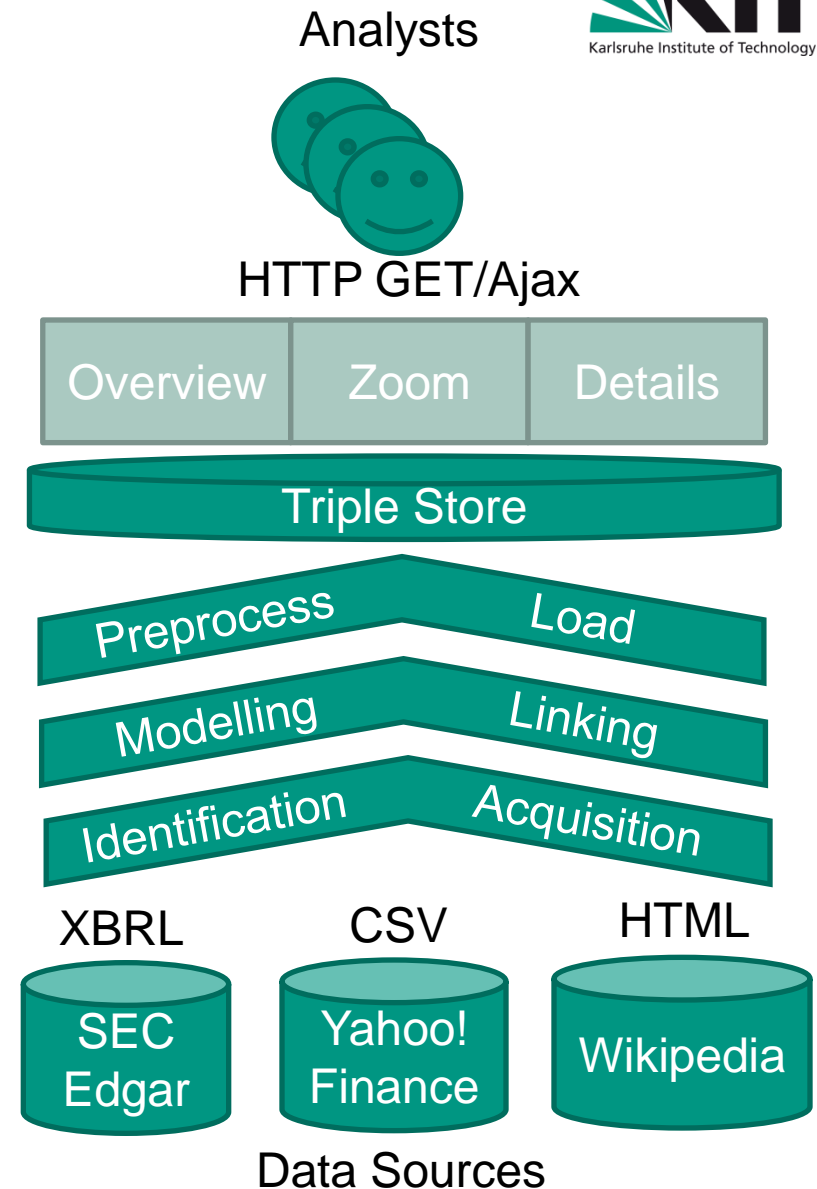


FIOS – How to allow analyses over finance data in RDF?

How to make analysts trust the data?

Shneiderman's Information Seeking Mantra

- Overview first
- Zoom and filter
- Details on demand



FIOS – Overview First

Analysts



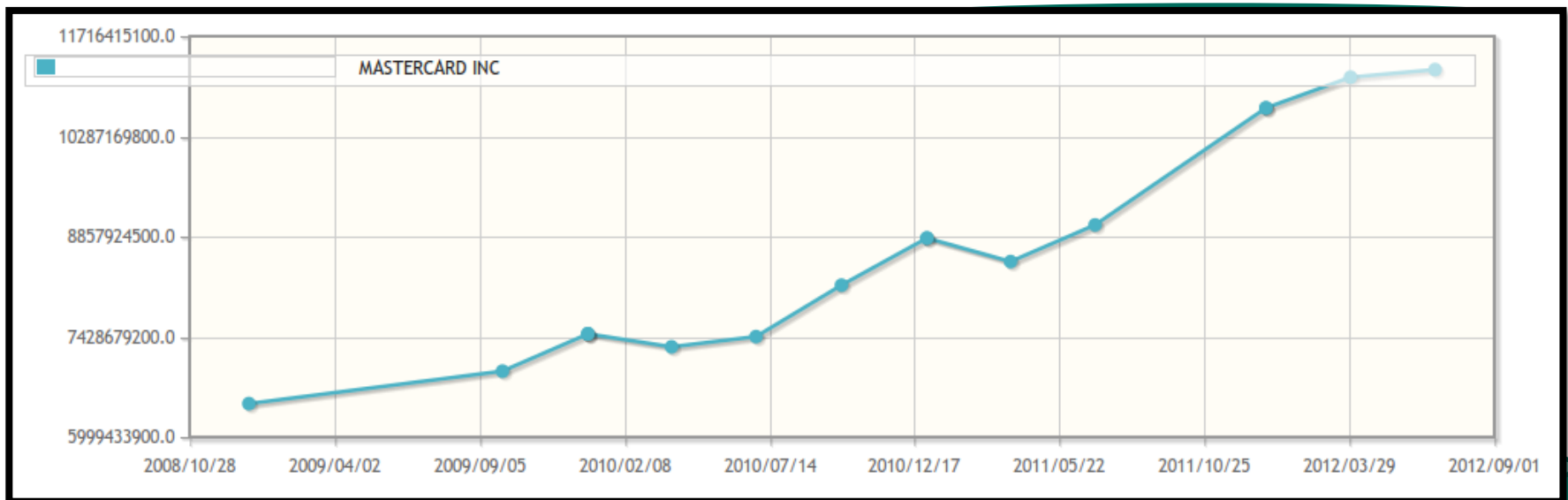
HTTP GET/Ajax

■ Total Assets for MASTERCARD INC

SPARQL
templates

Zoom

Details



SEC
Edgar

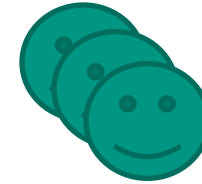
Yahoo!
Finance

Wikipedia

Data Sources

FIOS – Overview First

Analysts



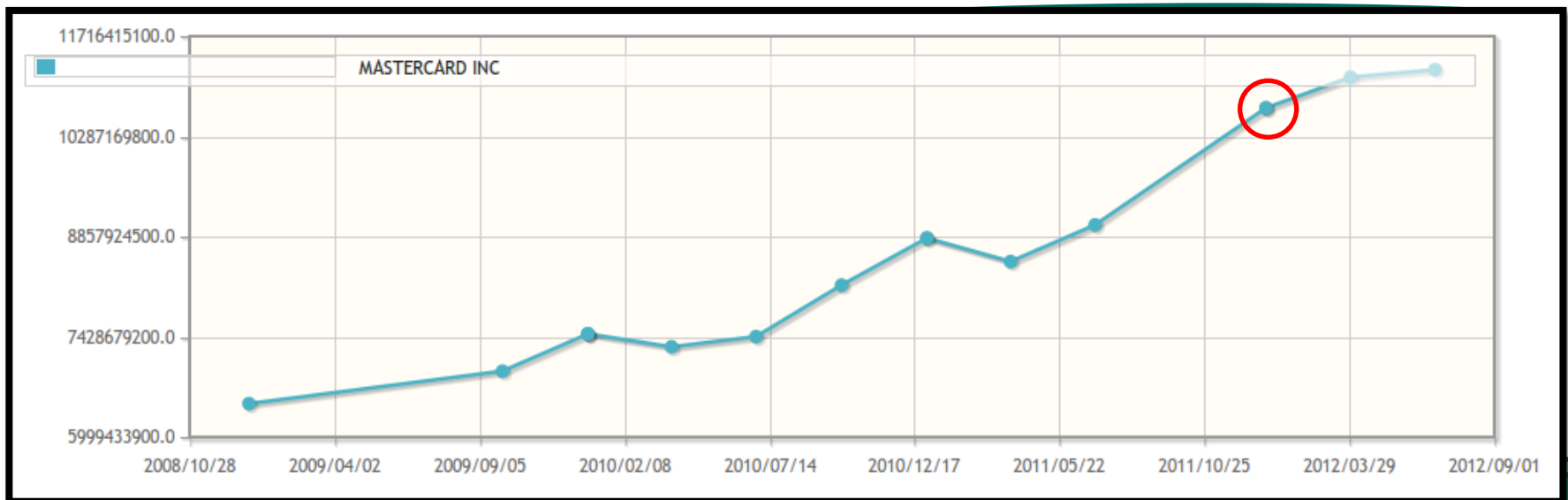
HTTP GET/Ajax

■ Total Assets for MASTERCARD INC

SPARQL
templates

Zoom

Details




SEC
Edgar


Yahoo!
Finance

Wikipedia









Data Sources

FIOS – Zoom and Filter



Cubes 



FIOS 2.0 Data Cube 



Dimensions



- ▼  **Date**
 -  Date
- ▼  **Issuer**
 -  SIC Level
 -  Company
- ▶  **Segment**
- ▼  **subject**
 -  Subject

Measures

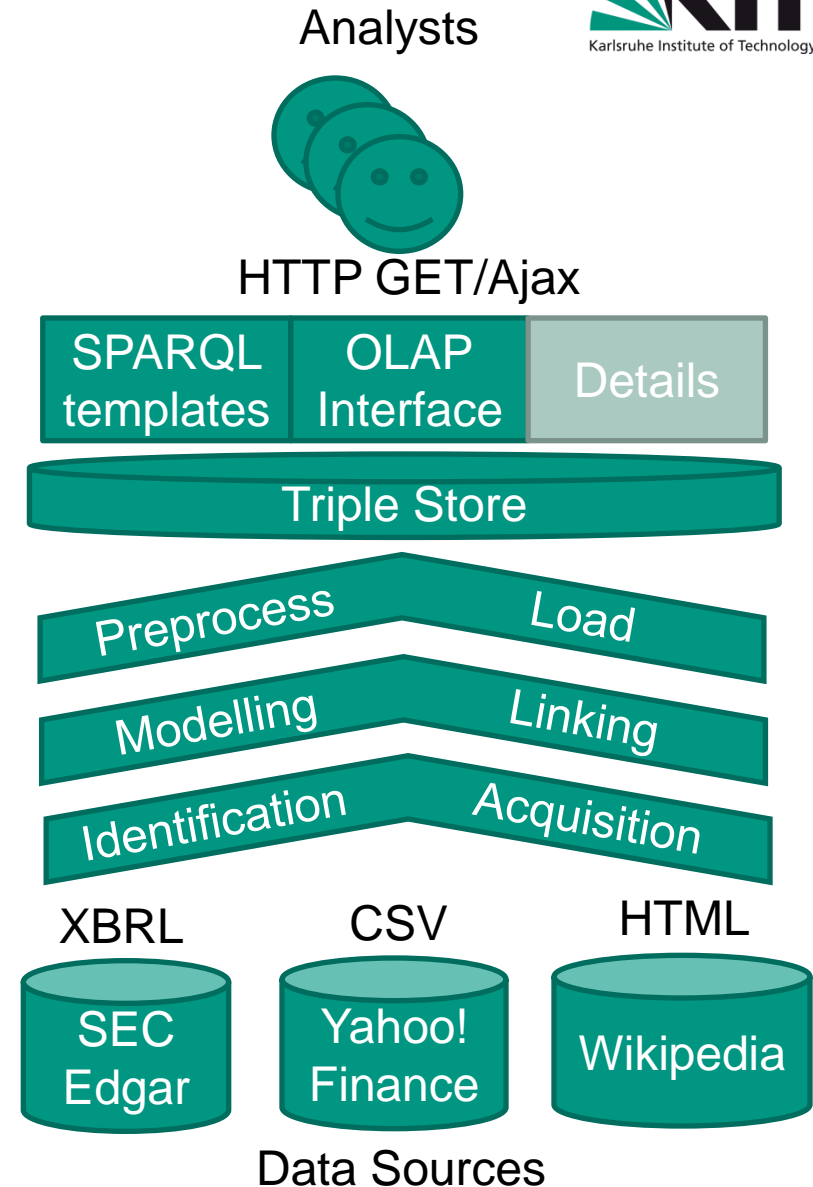
- ▼  **Measures**
 -  Obs value

Columns  Company Level 

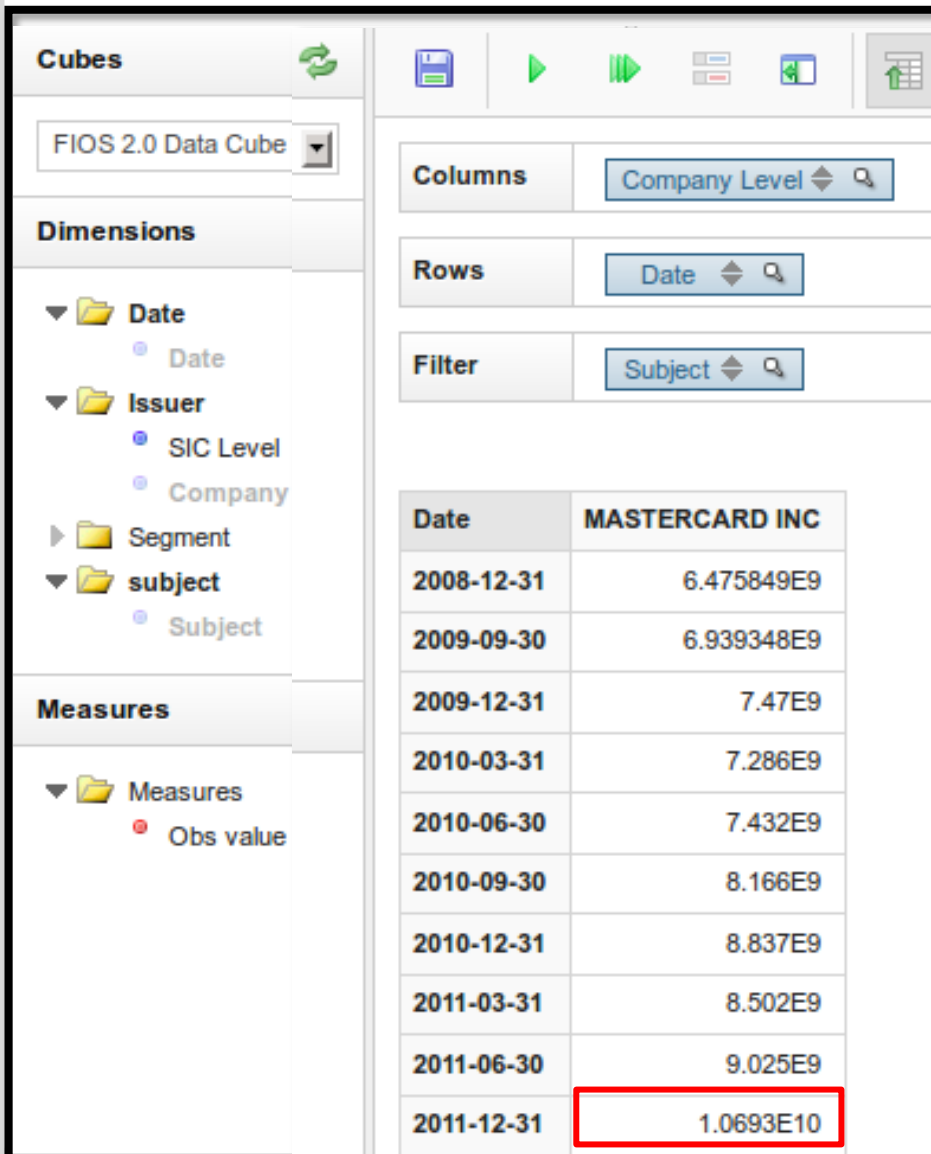
Rows  Date 

Filter  Subject 

Date	MASTERCARD INC
2008-12-31	6.475849E9
2009-09-30	6.939348E9
2009-12-31	7.47E9
2010-03-31	7.286E9
2010-06-30	7.432E9
2010-09-30	8.166E9
2010-12-31	8.837E9
2011-03-31	8.502E9
2011-06-30	9.025E9
2011-12-31	1.0693E10

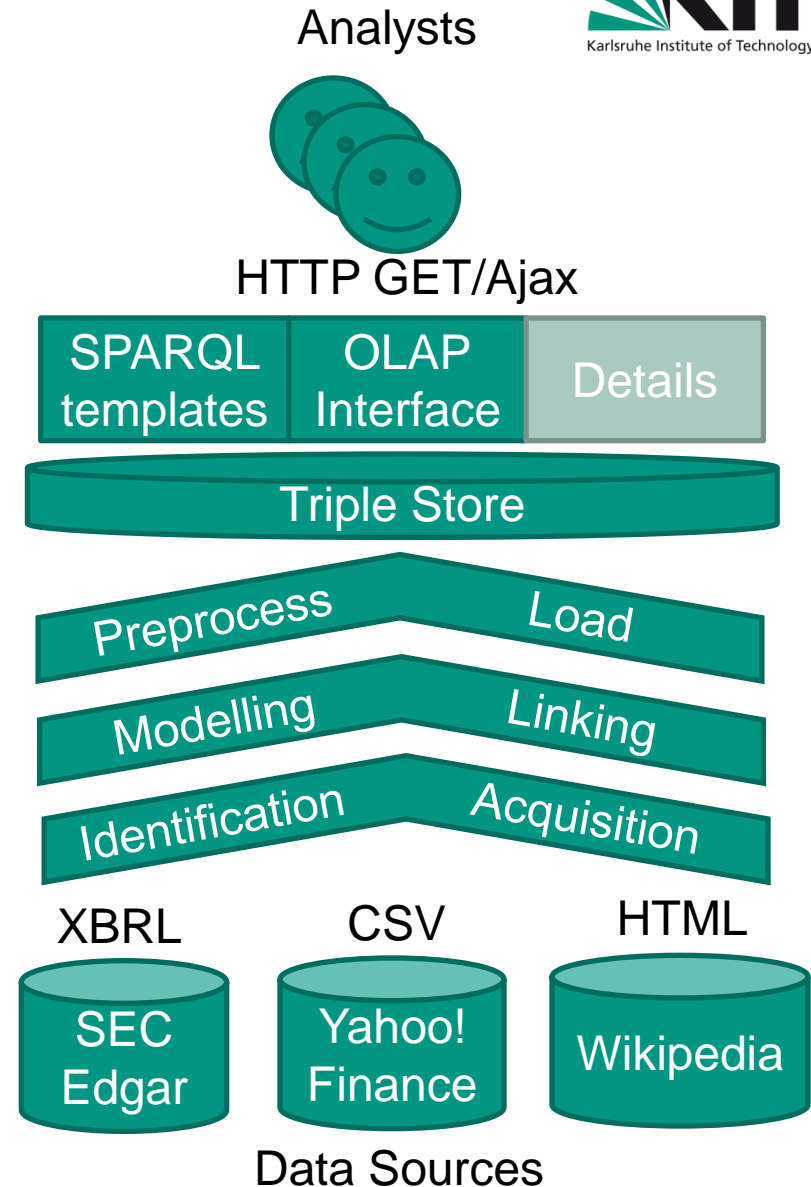


FIOS – Zoom and Filter



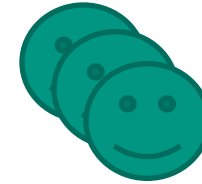
The screenshot shows the FIOS 2.0 Data Cube interface. On the left, there are panels for 'Cubes' (showing 'FIOS 2.0 Data Cube'), 'Dimensions' (with a tree view including Date, Issuer, Segment, and subject), and 'Measures' (showing 'Obs value'). The main area displays a table with columns 'Date' and 'MASTERCARD INC'. The table contains 11 rows of data, with the last row (2011-12-31) highlighted with a red box, showing a value of 1.0693E10. Above the table are controls for 'Columns' (Company Level), 'Rows' (Date), and 'Filter' (Subject).

Date	MASTERCARD INC
2008-12-31	6.475849E9
2009-09-30	6.939348E9
2009-12-31	7.47E9
2010-03-31	7.286E9
2010-06-30	7.432E9
2010-09-30	8.166E9
2010-12-31	8.837E9
2011-03-31	8.502E9
2011-06-30	9.025E9
2011-12-31	1.0693E10



FIOS – Details on demand

Analysts



HTTP GET/Ajax

SPARQL
templates

OLAP
Interface

LD
Browser

Anonymous Resource #285

Property	Value
qb:dataSet	■ http://public.b-kaempgen.de:8080/pubby/archive/1141391/0001141391-12-000006%23id
dcterms:date	■ 2011-12-31 (xsd:date)
?:issuer	■ http://public.b-kaempgen.de:8080/pubby/cik/1141391%23id
sdmx-measure:obsValue	■ 10693000000 (xsd:double)
is rdfs:seeAlso of	■ http://public.b-kaempgen.de:8080/pubby/archive/1141391/0001141391-12-000006%23id
?:segment	■ 0001141391 2011-12-31
?:subject	■ http://public.b-kaempgen.de:8080/pubby/vocab/us-gaap-2011-01-31%23Assets
rdf:type	■ qb:Observation

SEC
Edgar

Yahoo!
Finance

Wikipedia

Data Sources

FIOS – Details on demand

Analysts



HTTP GET/Ajax

SPARQL
templates

OLAP
Interface

LD
Browser

Anonymous Resource #285

Property	Value
qb:dataSet	■ < http://public.b-kaempgen.de:8080/pubby/archive/1141391/0001141391-12-0000069 >
dcterms:date	■ 2011-12-31 (xsd:date)
?:issuer	■ < http://public.b-kaempgen.de:8080/pubby/cik/1141391%23id >
sdmx-measure:obsValue	■ 10693000000 (xsd:double)
is rdfs:seeAlso of	■ < http://public.b-kaempgen.de:8080/pubby/archive/1141391/0001141391-12-0000069 >
?:segment	■ 0001141391 2011-12-31
?:subject	■ < http://public.b-kaempgen.de:8080/pubby/vocab/us-gaap-2011-01-31%23Assets >
rdf:type	■ qb:Observation

SEC
Edgar

Yahoo!
Finance

Wikipedia

Data Sources

Outline

- Mechanics in Financial Analysis
- XBRL – Challenges to Data Integration
- Financial Information Observation System (FIOS)
- **FIOS in the XBRL Challenge**
- Benefits of Semantics
- FIOS Competitors

FIOS in the XBRL Challenge – Setup

■ XBRL Challenge

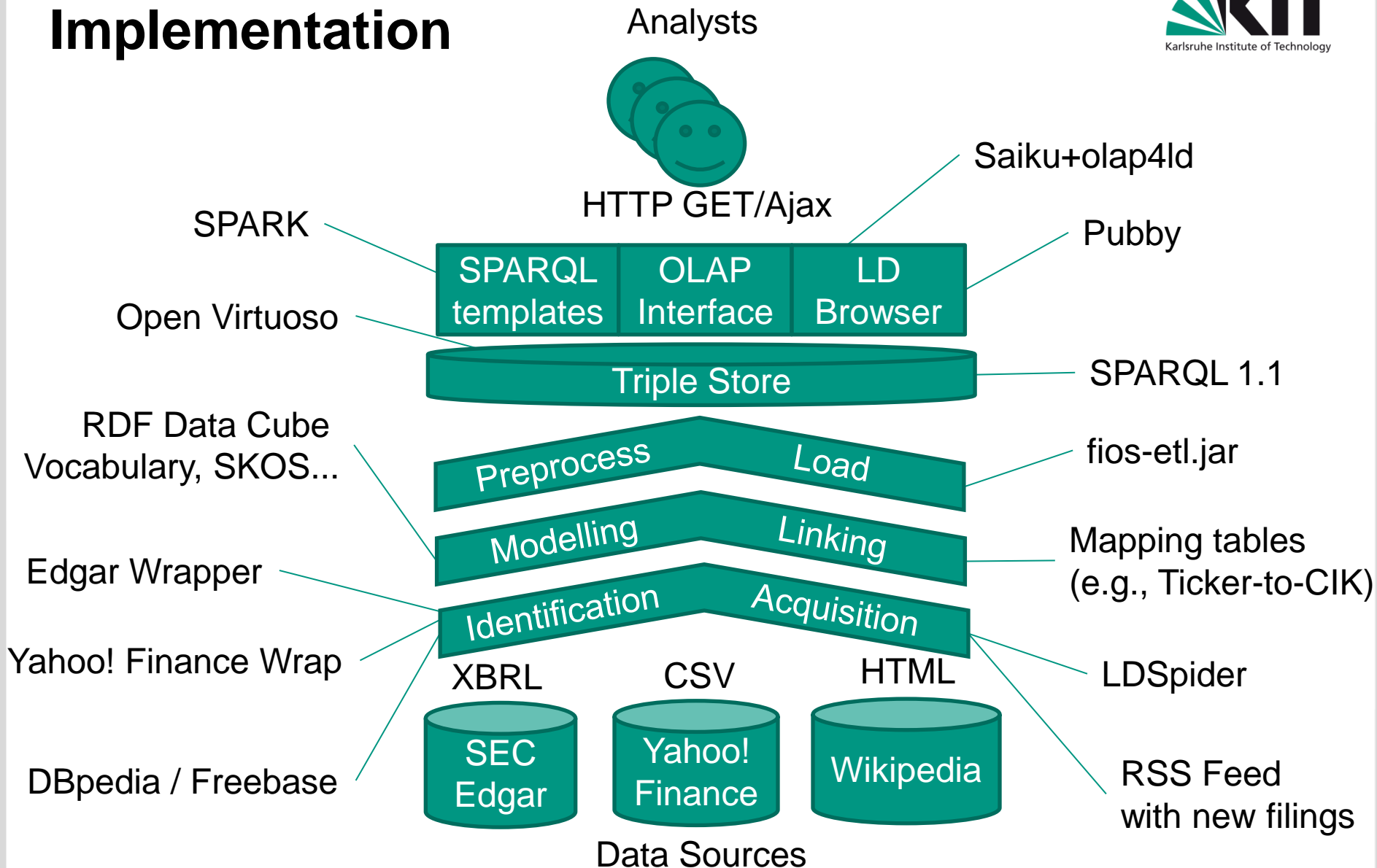
- 64 companies (finance, insurance and real estate)
- 3,781 financial concepts
- On average ETL components took 25min

■ Setup

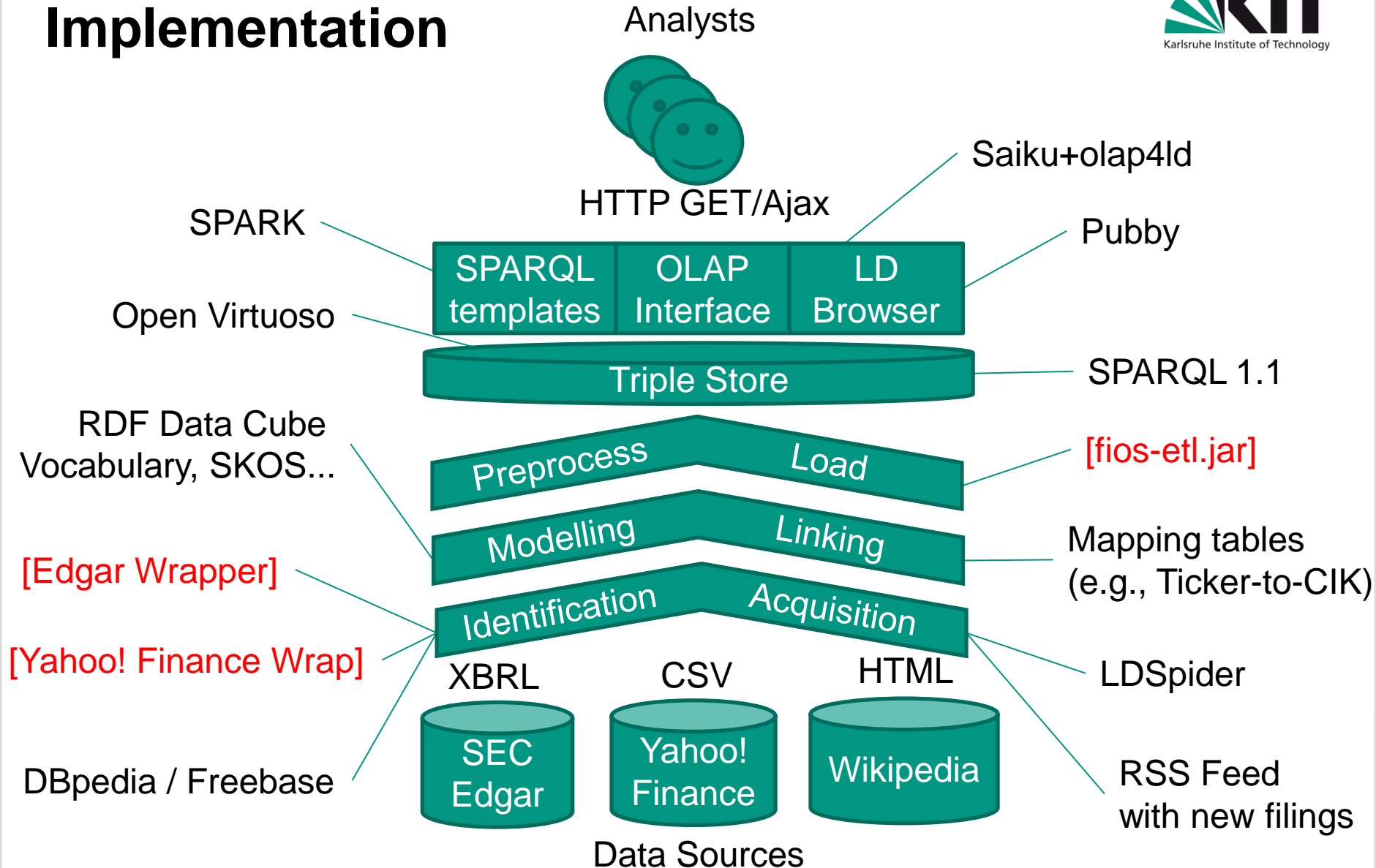
- 2673.330 MHz CPU and 1GB memory

[FIOS] http://fios.linked-data-cubes.org/FIOS_2_0/Queries/

FIOS in the XBRL Challenge – Implementation



FIOS in the XBRL Challenge – Implementation



Outline

- Mechanics in Financial Analysis
- XBRL – Challenges to Data Integration
- Financial Information Observation System (FIOS)
- FIOS in the XBRL Challenge
- **Benefits of Semantics**
- FIOS Competitors

Benefits of Semantics

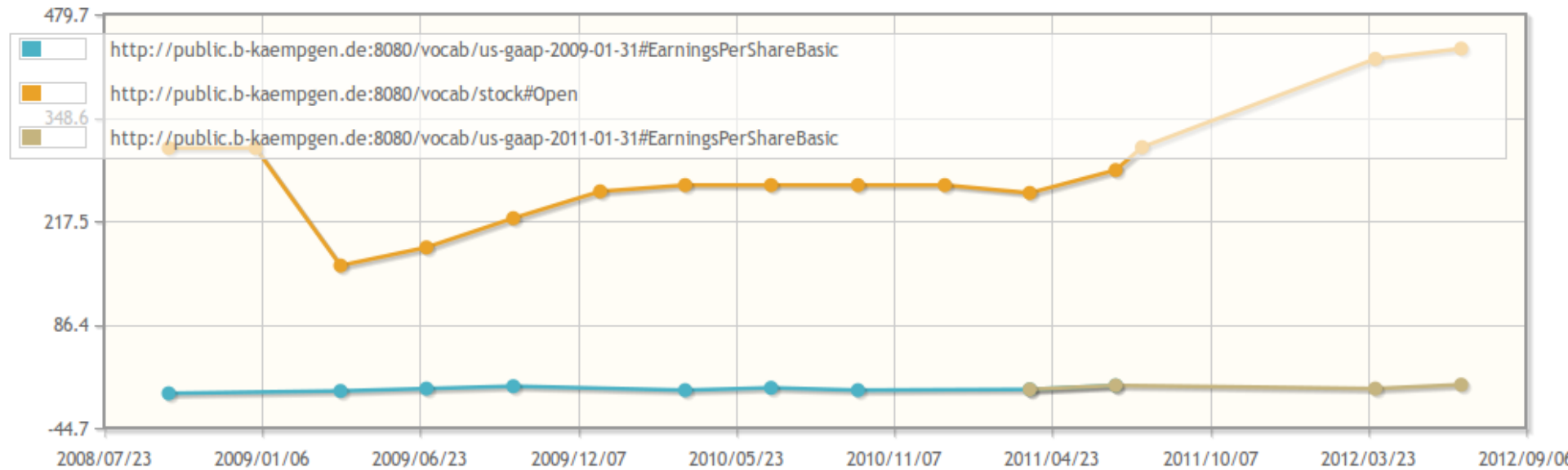
Task	Benefit
1) Identification & Acquisition	Modular access mechanisms
2) Modelling & Linking	Unified schema, vocabularies (QB/OWL)
3) Preprocessing & Loading	Well-defined equivalence semantics
4) Analysis	SPARQL 1.1, linking between interfaces

Benefits of Semantics

Task	Benefit
1) Identification & Acquisition	Modular access mechanisms
2) Modelling & Linking	Unified schema, vocabularies (QB/OWL)
3) Preprocessing & Loading	Well-defined equivalence semantics
4) Analysis	SPARQL 1.1, linking between interfaces

Benefits of Semantics

Task	Benefit
1) Identification & Acquisition	Modular access mechanisms
2) Modelling & Linking	Unified schema, vocabularies (QB/OWL)
3) Preprocessing & Loading	Well-defined equivalence semantics
4) Analysis	SPARQL 1.1, linking between interfaces



Earnings per Share versus Price per share for MASTERCARD INC

Benefits of Semantics

Task	Benefit
1) Identification & Acquisition	Modular access mechanisms
2) Modelling & Linking	Unified schema, vocabularies (QB/OWL)
3) Preprocessing & Loading	Well-defined equivalence semantics
4) Analysis	SPARQL 1.1, linking between interfaces

Task	Benefit
1) Identification & Acquisition	Modular access mechanisms
2) Modelling & Linking	Unified schema, vocabularies (QB/OWL)
3) Preprocessing & Loading	Well-defined equivalence semantics
4) Analysis	SPARQL 1.1, linking between interfaces

■ Opportunity: Continuously adding new sources

Challenge	Example
Seamless integration of interfaces and (new) data	Aggregated / densely displayed entities
Increase quality with new data	Ticker-CIK mapping evolves over time
Scale with new data	Analytical queries over many facts, Regularly loading of data

Outline

- Mechanics in Financial Analysis
- XBRL – Challenges to Data Integration
- Financial Information Observation System (FIOS)
- FIOS in the XBRL Challenge
- Benefits of Semantics
- **FIOS Competitors**

FIOS Competitors

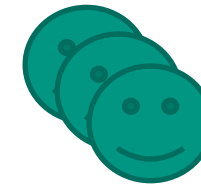
Work\Source	XBRL	Stock Quotes	Background Info	Focus
Calcbench / Sector3	++	-	-	Robustness, Search, Update intervals, Excel integration
BIXL / Midas	+	-	+	Semi-structured sources, multi-linguality
FIOS	+	+	+	Integration of structured data sources

Conclusions

- XBRL promising for Finance and SW people
- SW ready to help analysts with mechanics
- Opportunity to continuously add data sources

Thanks!

Analysts



HTTP GET/Ajax

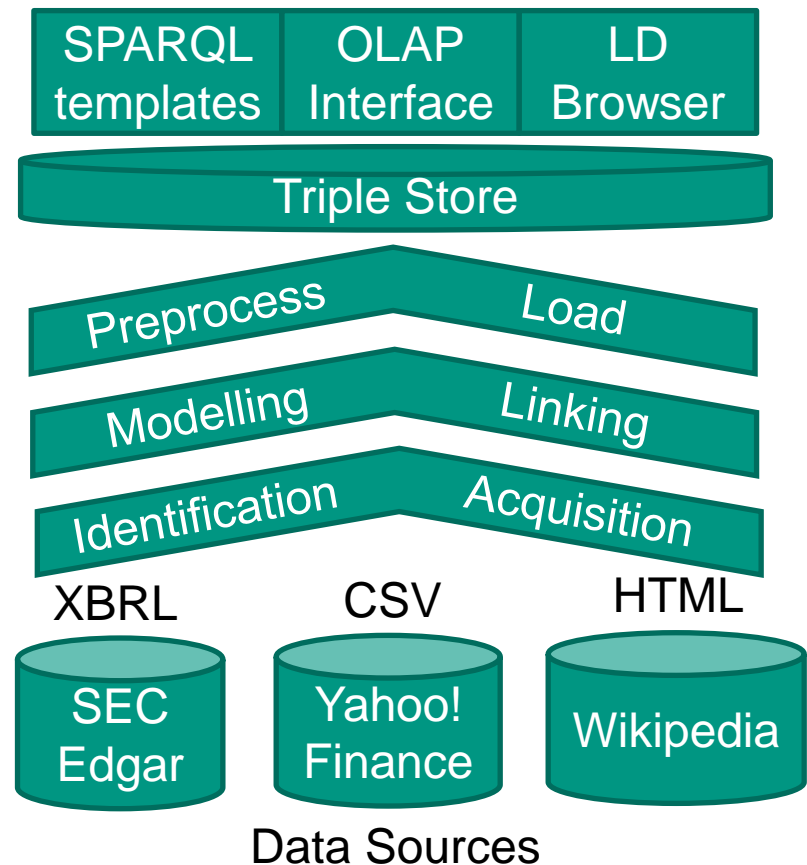
Benefit of Semantics

Modular access mechanisms

Unified schema, vocabularies

Well-defined equivalence semantics

SPARQL 1.1, linking between interfaces



References

- [XBRL SPEC] <http://www.xbrl.org/Specification/XBRL-2.1/REC-2003-12-31/XBRL-2.1-REC-2003-12-31+corrected-errata-2013-02-20.html>
- [XBRL example] <http://www.sec.gov/Archives/edgar/data/52827/000119312510238973/0001193125-10-238973-index.htm>
- [Hogan 2011] Hogan, A., Harth, A., Umbrich, J., Kinsella, S., Polleres, A., Decker, S.: Searching and browsing Linked Data with SWSE: The Semantic Web Search Engine. Web Semantics: Science, Services and Agents on theWorld WideWeb 9, 365–401 (2011)
- [KH2013] Kämpgen, B., Harth, A.: No Size Fits All - Running the Star Schema Benchmark with SPARQL and RDF Aggregate Views. In: ESWC (2013)
- [Yahoo! Finance Wrap] <http://yahoofinancewrap.appspot.com/>
- [Edgar Wrapper] <http://edgarwrap.ontologycentral.com/>
- [fios-etl.jar] <https://code.google.com/p/fios-etl/>